JOURNAL OF APPLIED PHYSICS

JOURNAL OF APPLIED PHYSICS

Vol. 24, No. 5 MAY, 1953

In This Issue

		Page
Arc Drop and Deionization Time in Inert Gas Thyratrons	S. Paksweer and R. Meyer	501
Pore Distribution in Porous Media	H. I. Meyer	510
Evaporation of Germanium Films from a Carbon Crucible	K. Lehovec, J. Rosen, A. MacDonald, and J. Broder	513
Propagation of Acoustic Waves in a Liquid-Filled Cylindrical Hole, Surrounded by an Elastic Solid	Edward V. Somers	515
Subharmonic Oscillations in Nonlinear Systems	Chihiro Hayashi	521
Null Characteristics of the Rotating Adcock Antenna System	James G. Holbrook	530
Method for the Quantitative Determination of Preferred Orientation	L. K. Jetter and B. S. Borie, Jr.	352
Free Space Radiation Impedance of Rhombic Antenna	Jesse Gerald Chaney	536
Effect of a Soller Slit on the Diffraction of X-Rays by Deformed Crystals	B. D. Cullity and Carl A. Julien	541
Current Element Near the Edge of a Conducting Half-Plane	Roger F. Harrington	547
Radiation Field of a Conical Helix	J. S. Chatterjee	550
Effect of Pressure on the Tensile Properties of Several Metals and Other Materials	P. W. Bridgman	560
Field Emitter: Fabrication, Electron Microscopy, and Electric Field Calculations	W. P. Dyke, J. K. Trolan, W. W. Dolan, and George Barnes	570
Elasto-Plastic Stress-Optical Effect in Silver Chloride Single Crystals	L. E. Goodman and J. G. Sutherland	577
Gas Flow Past Slender Bodies	Max M. Munk	584
Space Charge Spread of Reflected Electron Beams Studied by a Photographic Method	J. T. Wallmark	590
Metal Capillary Cathodes	H. Katz	597
Modes in Wave Guides Containing Ferrites	M. L. Kales	604
Response of Linear Time-Dependent Systems to Random Inputs	D. B. Duncan	609
Temperature Dependence of Low Frequency Fluctuations in Thermionic Emitters	T. B. Tomlinson	611
X-Ray Shadow Microscope	V. E. Cosslett and W. C. Nixon	616
Internal Friction and Young's Modulus of Cold-Worked Copper Single Crystals	J. Weertman and J. S. Koehler	624
Double Activators in Oxide Cathode Base Alloys	A. Eisenstein, H. John, and J. H. Affleck	631
Spectral Emissivities of Graphite and Carbon	R. J. Thorn and O. C. Simpson	633
Axial Motion of Electrons in a Linear Accelerator with $\beta=1$	James Swihart and Edward Akeley	640
Electric Permittivity of a Dilute Suspension of Membrane-Covered Ellipsoids	Hugo Fricke	644
Radiation Resistance of a Small Circular Loop in the Presence of a Conducting Ground	James R. Wait	646
Dynamic Mechanical Properties of Polyisobutylene	Edwin R. Fitzgerald, Lester D. Grandine, Jr., and John D. Ferry	650
Calibration of Photographic Emulsions for Electron Diffraction Investigations	L. S. Bartell and L. O. Brockway	656
Letters to the Editor:		
Titanium Replica for Electron Microscopy	Shigeto Yamaguchi and Tadayuki Nakayama	658
Effect of Internal Strains on Linear Expansion, X-Ray Lattice Constant, and Density of Crystals	Chester R. Berry	658
Reflection Coefficients for Certain Rough Surfaces	V. Twersky	659
Growth of Cadmium Sulfide Crystals	M. E. Bishop and S. H. Liebson	660
Comments on Louis A. Pipes' Letter	H. S. Kirschbaum	661
Correction to "A Mathematical Analysis of Parallel-Connected Magnetic Amplifiers with Resistive Loads"	Louis A. Pipes	661
Method of Preparing Replicas for Electron Microscopic Examinations	W. S. Smith and W. G. Kirchgessner	662
Fourier Representation of Buerger's Image-Seeking Minimum Function	William J. Taylor	662
R-PAC: Analog Recorder-Playback Computer for Crystal Analysis	Ray Pepinsky and Paul Jarmotz	663
Announcement		649
Index to Advertisers		xxviii

ELMER HUTCHISSON, *Editor*

EARLE C. GREGG, *Assistant Editor*

Board of Editors

RAY E. BOLZ
LYLE D. BORST
WALLACE J. ECKERT
ROBERT D. HEIDENREICH
LEONARD B. LOEB
W. JAMES LYONS
LOUIS A. PIPES
ROMAN SMOLUCHOWSKI
W. R. SMYTHE

Publication of the American Institute of Physics

OFFICERS

GEORGE R. HARRISON, *Chairman*
GEORGE B. PEGRAM, *Treasurer*

GOVERNING BOARD

J. W. BUCHTA
KARL K. DARROW
G. J. DIENES
S. A. GOUDSMIT
GEORGE R. HARRISON
FREDERICK V. HUNT
DEANE B. JUDD
HUGH S. KNOWLES
C. C. LAURITSEN
WILLIAM F. MEGGERS
PHILIP M. MORSE
BRIAN O'BRIEN
GEORGE B. PEGRAM
I. I. RABI
DUANE ROLLER
FREDERICK SEITZ
WILLIAM SHOCKLEY
RICHARD M. SUTTON
J. H. VAN VLECK
MARK W. ZEMANSKY

Consultant on Printing

MELVIN LOOS

ADMINISTRATIVE STAFF

HENRY A. BARTON, *Director*
WALLACE WATERFALL, *Executive Secretary*
THEODORE VORBURGER, *Advertising Manager*
RUTH F. BRYANS, *Publication Manager*
BETTY H. GOODFRIEND, *Assistant Secretary*
EILEEN B. NEUBERGER, *Circulation Manager*
KATHRYN SETZE *Assistant Treasurer*

The JOURNAL OF APPLIED PHYSICS, published monthly at Prince and Lemon Sts., Lancaster, Pa., is devoted to physics in its role as the science basic to other natural sciences and to the arts and industries. Previous to 1937 this publication was known as PHYSICS. The JOURNAL OF APPLIED PHYSICS publishes editorials and reviews, as well as technical papers of applied physics. Articles appearing in it are indexed regularly, according to their field, by Physics Abstracts, Chemical Abstracts, Engineering Index, Inc., Industrial Arts Index, etc.

Manuscripts should be submitted to Elmer Hutchisson, Editor, Case Institute of Technology, Cleveland 6, Ohio. Unless otherwise stated, submission of a manuscript is a representation that it has been neither copyrighted, published, nor submitted for publication elsewhere. The authors' institutions are requested to pay a **publication charge** of $8 per page which, if honored, entitles them to 100 reprints without further charge. Instructions will be sent with galley proofs.

Subscriptions, renewals, and orders for back numbers should be addressed to the American Institute of Physics, 57 East 55 Street, New York 22, New York.

Subscription Price	U.S. and Canada	Elsewhere
To members of the American Institute of Physics and Affiliated Societies	$10.00	$11.00
To all others	12.00	13.00

Back Number Prices

Yearly back number rate when complete year is available: $14.00. Single copies: $1.25 each.

Changes of address should be sent to the American Institute of Physics; **corrected proofs** should be sent to the Publication Manager, American Institute of Physics. **Advertising rates** supplied on request. Orders, advertising copy, and cuts should be sent to the American Institute of Physics.

Entered as second class matter January 22, 1937, at the Post Office at Lancaster, Pennsylvania, under the Act of March 3, 1879. Accepted for mailing at the special rate of postage provided for in paragraph (d-2), section 34.40, P. L. & R. of 1948, authorized May 2, 1932.

Journal
of
Applied Physics

Volume 24, Number 5 May, 1953

Arc Drop and Deionization Time in Inert Gas Thyratrons

S. Pakswer and R. Meyer
Research Department, The Rauland Corporation, Chicago 41, Illinois
(Received September 5, 1952)

The energy balance is discussed and a semi-empirical equation for recovery time is derived; the discussion is based on diffusion alone without consideration of cumulative ionization and metastable states. Corroborating the theoretical considerations, measurements were made of arc drop and recovery time in thyratrons filled with mixtures of Xe with Ar, Ne, and He. The spectral emission lines in these discharges were photographed in the visible range. It was found that the arc drop remained fairly constant down to low concentrations of Xe and then increased steeply. Lines of the gas other than Xe also appeared only at low concentrations of Xe. The recovery time is a monotone declining function of the mixture of the heavy and the light components. The spectral lines are tabulated.

TWO requirements received particular attention in the design of low voltage gas-filled thyratrons for inverter applications. It proved necessary to develop tubes with the lowest possible arc drop and the shortest possible deionization time. As it was felt that little information was available about the over-all energy balance in and the disintegration of a thyratron plasma, measurements were made of arc drop, recovery time, and spectral characteristics in mixtures of gases in which the activation and ionization potentials and the atomic weights of the components varied considerably. Some observations of such conditions had been previously disclosed by Edwards and Smith.[1]

EXPERIMENTAL SET-UP

Most studies of thyratrons, as described in the literature, have been based on completely finished, sealed-off tubes. Such an approach has 3 drawbacks: (1) many tubes are necessary to obtain data, (2) the data are subject to variations due to variations in geometry and processing of individual tubes, and (3) the pressure in a sealed-off envelope is undeterminable because of unknown temperature conditions of the gas. It was decided therefore to make all tests on processed single tubes, while still attached to the pump manifold. Thus, large series of data could be obtained in a comparatively short time. The pressures were measured by means of a McLeod gauge included in the system, and neither effects resulting from pressure difference in a tube running at around 200°C nor effects resulting from thermal diffusion were observed.

All experiments were made on a slightly modified version of a C5F14 tube. This tube had an indirectly heated cathode, coated with a mixture of Ba and Sr oxides, a tantalum anode, and a single cylindrical hole in the control grid having a diameter of 0.312 in. and a height of 0.244 in. The bottom of the hole is $\frac{5}{32}$ in. away from the hole in the cathode baffle, and the top of the hole is $\frac{1}{4}$ in. distant from the bottom of the anode.

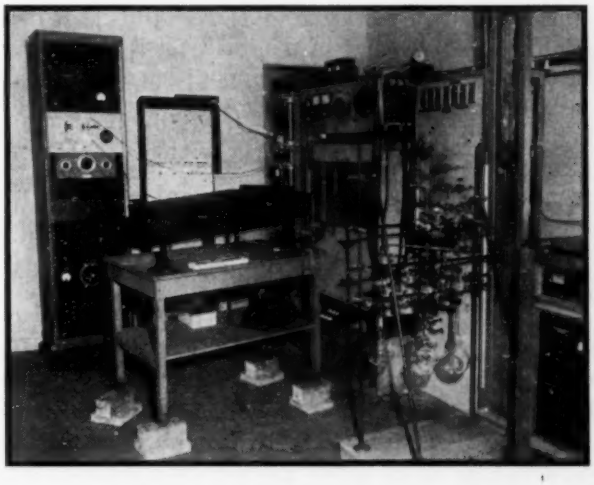


FIG. 1. Experimental set-up.

[1] U. S. Patent 2,567,369.

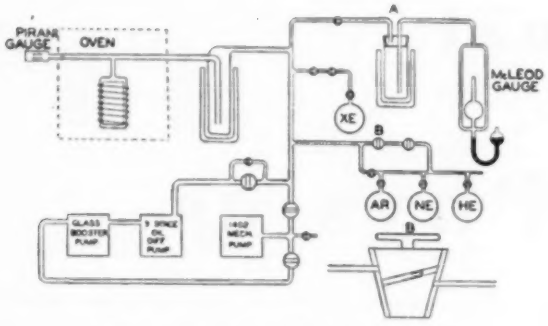


FIG. 2. Outline of exhaust system.

The experimental apparatus is shown in Fig. 1. The cabinet to the left contains the deionization tester. A schematic drawing of the glass manifold and the pump system is shown in Fig. 2. After some experimentation it was found that the best way of obtaining a mixture of xenon and any other gas was to fill the manifold with Xe at a pressure corresponding to its percentage in the desired mixture and then slowly add the additional gas to the desired total pressure. For this purpose, a small glass tube drawn down at one end to an opening approximately 0.001 in. in diameter, was cemented into the bore of stopcock *B*, through which the gas confined in the trap at nearly atmospheric pressure could escape very slowly. Another special feature of the system was the trap *A*, containing metallic Na and held at dry ice temperature, which prevented Hg vapor from the McLeod gauge from diffusing into the manifold.

For arc drop and starting voltage measurements a regulated dc power supply was used. The grid was connected to the anode by a 10 000-ohm resistor, the tube thus operating as a diode and not as a thyratron. The circuit of the deionization tester is shown in Fig. 3. The amplified sine wave signal of a Hewlett-Packard audio-frequency oscillator is applied to the grid of a triode operating in a class *B* circuit. The modulated output is transmitted to the anode of the thyratron under test, the grid of which is connected by means of a switch to either a -45-v or to a $+22.5$-v battery through a 10-K resistor. The frequency of the oscillator is raised until the tube continues firing upon application of -45 v. At this frequency the required recovery time equals the duration of the negative half-cycle, and, since this frequency can be read directly in cycles/sec on the frequency scale of the Hewlett-Packard oscillator, the recovery time can be expressed in microseconds using the formula

$$\text{recovery time } (\mu\text{sec}) = 10^6/2\times\text{cycles}.$$

Measurements of deionization time with sine wave instead of square pulses have two drawbacks: (1) some distortion of the sine wave is effected by class *B* operation, and (2) the tube starts to deionize when the sine wave voltage drops below the arc drop voltage, thus showing somewhat shorter deionization times than actually occur. The advantage of the sine wave method is its great simplicity.

The spectrograph used was a grating instrument made by the Central Scientific Company, catalog No. 87102. All photographs were made on Eastman super-sensitive panchromatic plates with 10–15 minute exposure.

ARC DROP

(a) Energy Balance

Basically, in a gas-filled rectifier tube, electrons are transported from a low work-function oxide-coated cathode to a high work-function electrode made, for example, out of tantalum. Were these two electrodes in direct contact and in equilibrium, a contact potential between the electrodes would be established corresponding to the difference of work functions. Assuming $\varphi_a = 4.1$ ev for Ta and $\varphi_{c'} \approx 1.0$ ev for the oxide-coated electrode and correcting $\varphi_{c'}$ for a field-enhanced emissivity due to a positive space charge of ~ 10 v at the cathode by the assumption of $J_e/J_{th} \approx 2.0$ to $\varphi_c = 0.93$, it can be seen that this contact potential would be 3.17 volts, with the tantalum electrode being charged negatively. The full cathode emission would only flow continuously into the Ta electrode if a corresponding positive voltage were applied to it; hence, 3.2 volts would be the minimum obtainable arc drop for this electrode combination. Using an anode of lower work function and a cathode of higher work function would reduce this voltage drop, as has been shown by G. Medicus and G. Wehner.[2] However, the actually observed arc drops in grid-controlled Xe rectifiers in continuous dc operation are of the order of 6.5 to 8.5 volts. The difference is due to an increase in the kinetic energy of the electrons to compensate for loss of electrons and ions in the plasma by ambipolar diffusion, recombination, and radiative processes in the positive column. Also, part of the energy is transferred to the anode by residual kinetic energy of the electrons striking it and appears as anode fall. Since $(V_a + \varphi_a)\, i = A \cdot a\sigma T^4$ [3] (V_a = anode fall, A = anode surface area, a = total emissivity of Ta = 0.45, σ = Stefan Boltzmann constant, T = temperature of the anode in °K), the anode fall can be roughly estimated from the anode temperature,

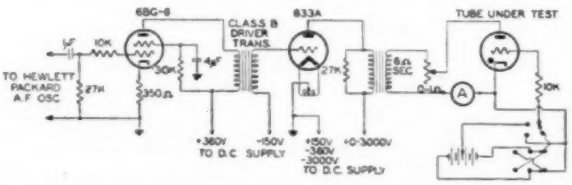


FIG. 3. Circuit diagram of the deionization time tester.

[2] G. Medicus and G. Wehner, J. Appl. Phys. **22**, 1389 (1951).
[3] Leonard B. Loeb, *Fundamental Processes of Electrical Discharge in Gases* (John Wiley & Sons, Inc., New York, 1939), p. 615.

neglecting convection and conduction losses),* to 0–3 volts. The difference between the actual arc drop and $\varphi_a - \varphi_c + V_a$ is lost in the plasma.

The cathode fall does not enter into the energy balance as long as dc arc drop is measured. This results from the fact that dc arc drop voltages are observed which are less than the ionization potentials, a fact which has been satisfactorily explained.[4] According to this explanation a sheath of positive ions appears at the cathode which raises the space potential in the cathode drop beyond the ionization potential, thus providing enough acceleration for electrons to ionize atoms in the positive column. But, since this sheath is continuously replenished by ions from a dc discharge, the energy spent in its formation can be omitted. However, should the arc drop be measured with ac, for example, by the watt-meter method,[5] this sheath would have to be reformed at each cycle with an initial starting arc drop of at least 12.1 volts (ionization potential of xenon).

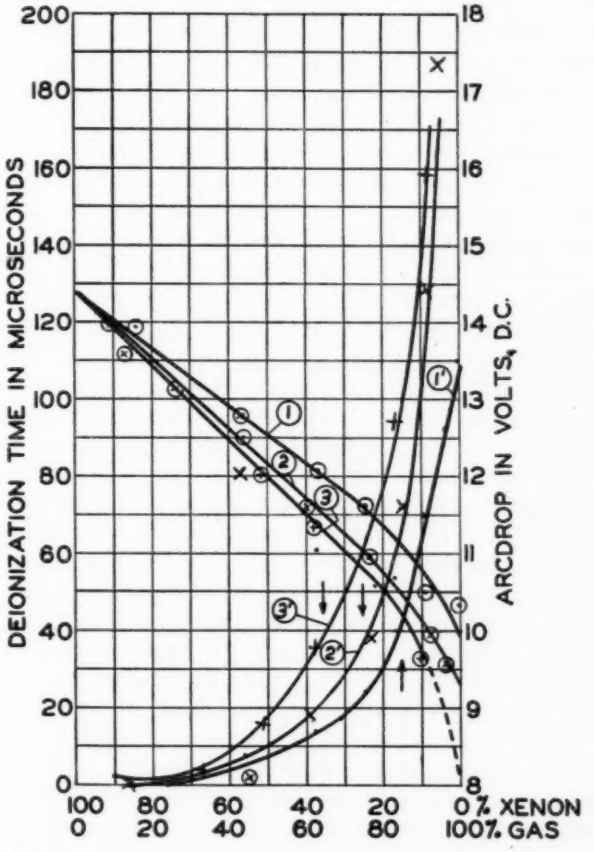


FIG. 4. Deionization time (Curves 1, 2, 3) and arcdrop (Curves 1′, 2′, 3′) in mixture of xenon with argon (Curves 1, 1′, and ·), neon (Curves 2, 2′, and ×), helium curves 3, 3′ and +). The test points of deionization measurements are enclosed in a circle. Pressure, 64 microns. Tube current, 3A.

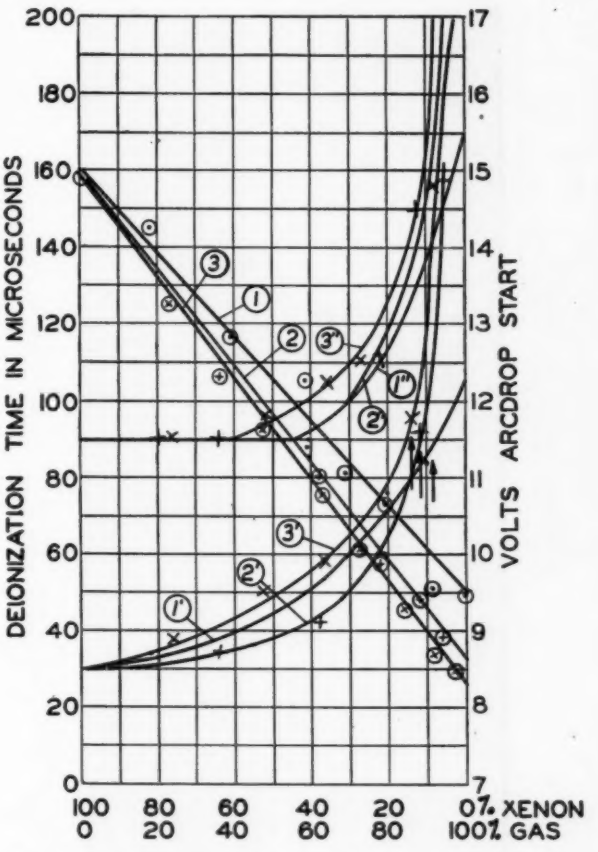


FIG. 5. Deionization time (Curves 1, 2, 3), arc drop (Curves 1′, 2′, 3′) and starting voltage (Curves 1″, 2″, 3″) in mixture of xenon with argon (Curves 1, 1′, 1″, and ·), neon (Curves 2, 2′, 2″, and +), helium (Curves 3, 3′, 3″, and ×). The test points of deionization measurements are enclosed in a circle. Pressure, 100 microns. $i_0 = 3A$.

Actually, the starting voltage of such tubes is 12–13 volts. Thus, the watt-meter method would show higher arc drops than the dc method, and the arc drop would depend on the peak forward voltage and the frequency of the wave. It should be mentioned at this point that additional complication can be expected in arcs in which the emission current of the cathode is higher than the tube current and an electronic instead of a positive ion space charge can be formed at the cathode.[6]

The energy losses in the plasma are the result of the radiation, recombination, and increase in kinetic energy of electrons necessary to compensate for the loss of electrons and ions by diffusion to sustain a constant current of the density $j = Ne\bar{V}$, where N is the number of charged particles and $\bar{V}$ their drift velocity.

The additional voltage which must be applied to compensate for diffusion losses can be determined by considering the current density with and without a diffusion term. Suppose, with diffusion,

$$j = ne\mu E - D_a e(dn/dx)$$

* Convection and conduction losses are counterbalanced by absorbed radiation from hot parts such as the grid and cathode, from radiating transitions in the gas and heat convection from the hot plasma.

[4] See reference 3, p. 636.

[5] E. K. Smith, Elec. Eng. 69, 419–422 (1950).

[6] D. E. Marshall, *Industrial Electronics Reference Book* by Electronic Engineers of the Westinghouse Corporation (John Wiley and Sons, Inc., New York, 1948), Chapter IV, p. 52.

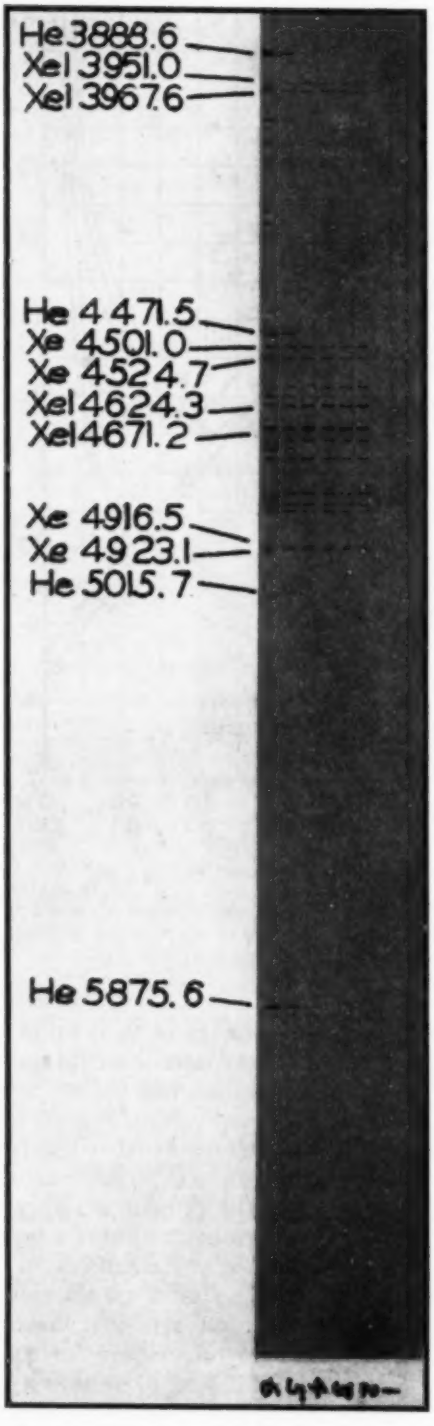


FIG. 6. Spectrogram of xenon-helium mixtures at 64-microns pressure, the consecutive spectrograms corresponding to: 1–10 percent, 2–33 percent, 3–42.5 percent, 4–62.2 percent, 5–85.3 percent, 6–93.7 percent helium.

(where D_a is the coefficient of ambipolar diffusion, n the number of charged particles, μ their mobility, and E the gradient in the positive column), and with no diffusion

$$j = ne\mu E_0.$$

If an equal current is maintained in the two cases,

$$ne\mu E_0 = ne\mu E - D_a e(dn/dx),$$

so that $\int (E - E_0)dx = D_a/\mu \int dn/n$. The expression on the left is the additional voltage which must be applied to overcome the diffusion loss, and

$$V_{\text{arc drop diffusion}} = (D_a/\mu) \ln n + \text{const} = 2/3 U_M \ln n + \text{const},$$

where U_M is the average energy of the carriers.[7] Since n, the number of charge carriers, increases with the current, the above expression is in accord with the experimentally observed fact that the arc drop increases with the current drawn through the tube. Also, if the supply of electrons from the cathode is not sufficient to carry all the current, additional kinetic energy must be imparted to the electrons and this results in a higher dc and ac arc drop.

(b) Arc Drop in Gas Mixtures

The above is true in cases of lowered tube pressure or dilution of gas with a gas of higher activation and ionization potential. The latter condition is illustrated in Figs. 4 and 5, showing measurements made on two different tubes employing mixtures of Xe with Ar, Ne, and He. It can be seen that with a diminishing concentration of Xe the arc drop remains just constant until, at still lower Xe percentages, the arc drop begins to rise rapidly to the arc drop of the pure additional gas. Spectrographic investigation of the discharge, as shown on a sample in Fig. 6 and in the appendix, shows that the main group of persistent Xe lines remains of equal intensity throughout the spectrum down to concentrations of 4–5 percent of Xe. Some additional Xe lines (mostly lines of ionized xenon) appear at higher arc drops in mixtures and grow progressively stronger. The lines of the additional gas are absent in the spectrum until the Xe content falls below 25 percent. The place where these lines first appear is marked in Figs. 4 and 5 by arrows on the arc-drop curves. From this point on these lines increase rapidly in intensity. Two observations need further explanation. (1) The spectral lines, especially of Ne and He, appear at arc-drop voltages which are well below the ionization (15.7; 21.5; 24.5 ev), excitation (11.5; 16.5; 20.91), or metastable (11.5; 16.6; 19.7) potentials of Ar, Ne, or He, respectively. This can be attributed to the previously described formation of an ion sheath at the cathode, containing a gradually increasing amount of the additional ions and thus imparting higher and higher velocity to some electrons as suggested by Loeb.[4] Thus, cumulative ionization cannot account for the appearance of lines in the dc arc, but it can be present at the start of the arc. To verify this result, some measurements of the dc start were plotted in Fig. 5. Since the measurements were not oscilloscopic, they do not give exact initial starting

[7] S. C. Brown, Elec. Eng. 71, 501–503 (1952).

volts, but indicate qualitatively the rapid rise of starting voltage with decreasing Xe content. (2) Spectrograms made on tube 2 (Fig. 5) showed for Ne and He that the spectral lines of the second element appeared at much lower Xe content than in tube 1 of Fig. 4 (for example, neon ~12 percent Xe against 25 percent Xe). Since tube 2 was run at higher filling pressure (100μ against 64μ), several spectrograms were made of this tube with neon at both pressures. These spectrograms are shown in Fig. 7, and it can be seen that at lower pressure spectral lines appear at a higher concentration of Xe than at higher pressure. This must be due, not to a difference in arc drops, but to a variation in starting voltage, thus again confirming the previous considerations.

DEIONIZATION TIME

The deionization time, or rather the "recovery time," of thyratrons ("Freiwerdezeit") is a function of many variables, such as tube geometry, negative grid voltage, grid resistance,† tube current prior to deionization, pressure, and nature of the gas. In making the tests, the first 3 variables were kept constant. The influence of tube pressure and tube current was measured in pure Xe and is shown in Figs. 8 and 9. In these figures, the crosses correspond to experimental data whereas the lines pertain to data calculated according to the experimental equation

$$t=(p+27)\times 0.309(\ln_e i+3.00361). \qquad (1)$$

In this equation t is in microseconds, p in microns, and i in amperes. It can be seen that t is a sum of t' and t'' where t' is a pressure-(diffusion) dependent and t'' a pressure-independent deionization. t'' can in this pressure region, where the mean free path is of the order of the diffusion length, be attributed to ions leaving the plasma without collisions, that is, to the transit time of the ions. It takes a Xe ion approximately 28.5 μsec to travel 1 cm at 700°K.

Keeping also p and i constant, measurements of recovery time were made on various gas mixtures. Edwards and Smith[1] observed that in xenon-argon mixtures the deionization time was cut roughly to one-half of the deionization time of each of the components, and this was attributed to the presence of metastable atoms. The present tests, as indicated in Figs. 4 and 5, show an almost linear decrease of recovery times from the heavier to the lighter component.

The dependence of deionization time on the atomic weight does not lend itself to calculation because, as will be shown later, several functions of m are involved, some of which are little known, and the balance of different factors can hardly be quantitatively established. Therefore, in Figs. 4 and 5 the drawn lines correspond to the most probable slopes of experimental data, and it can be seen that these lines unite points approximately in the inverse ratio of atomic weights if the atomic weight of a mixture is considered to be $fm_{Xe}+(1-f)m_g$ (where f is the fraction of Xe in the mixture).

An equation of the type used for the calculation of

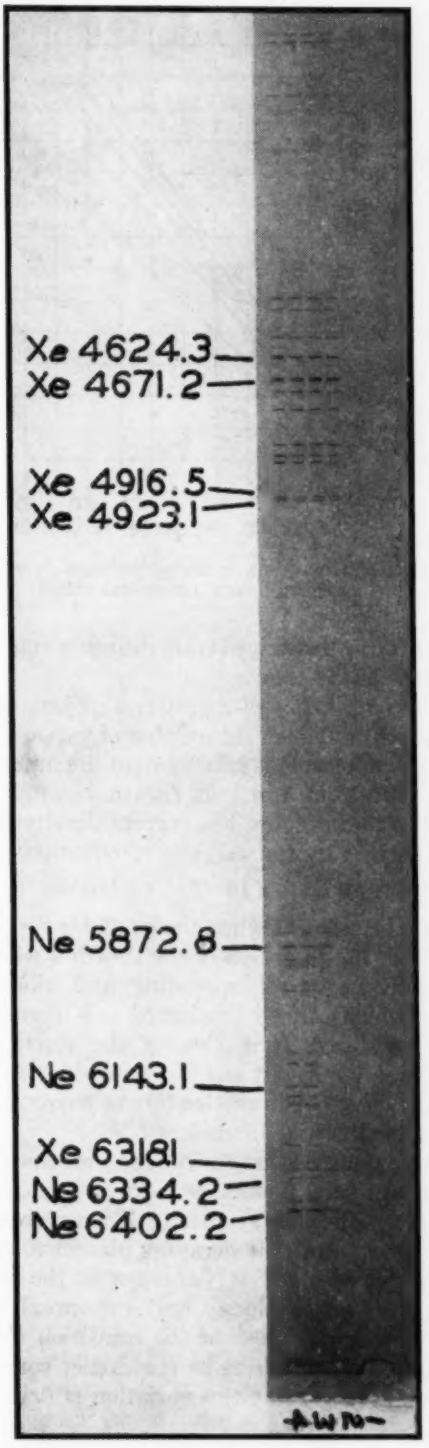


FIG. 7. Spectrogram of xenon-neon mixtures. 1—Neon 79 percent, pressure 98 microns. 2—Neon 79 percent, pressure 64 microns. 3—Neon 87 percent, pressure 103 microns. 4—Neon 87 percent, pressure 57 microns.

† For the effect of grid resistance and grid potential on recovery time see reference 10 and L. Malter and M. R. Boyd, Proc. Inst. Radio Engrs. 39, 636–643 (1951).

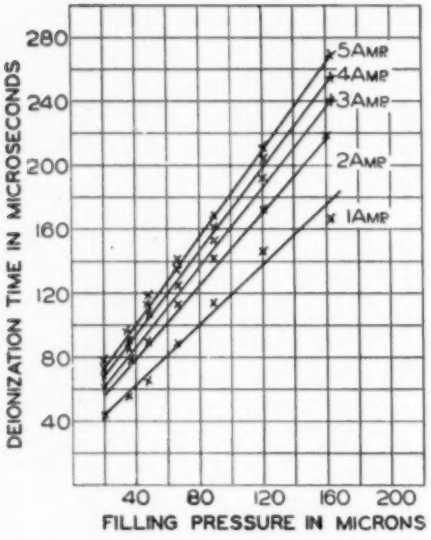


FIG. 8. Deionization time *vs* filling pressure in xenon. ×'s—measured points. Lines—calculated.

the points can be derived from diffusion considerations to a probe in the plasma.

The recovery time of a grid in a plasma occurs at a certain point at which the number of ions arriving from the decaying plasma is smaller than the number of ions flowing out from the grid sheath through the grid resistor to the cathode. The current density in the grid sheath is given by the well-known equation

$$j=2.336\times10^{-6}E^{\frac{1}{2}}/(1.824m)^{\frac{1}{2}}d^2. \qquad (2)$$

Thus, at the moment when the available supply of ions is depleted, the thickness of the sheath d will increase. This phenomenon of spreading and thickening of sheaths, which can be considered as a reversal of the process of sheath formation at the start, has been investigated by Koch[8] and Romanowitz.[9] A complete investigation of conditions leading to recovery has been made by Malter and Johnson.[10]

Let it be assumed that the tube regains control always at a constant value of sheath thickness (which is a function of m and the tube geometry). Then $j=\text{const}_1 f_1(m)$.

The current from the decaying plasma to the grid is given by $j\cdot A=n\cdot q\cdot\bar{v}\cdot A$ (where n is the number of charges, $\bar{v}$ their drift velocity, and A the area). $\bar{v}$ can be a function of time because of the transition from ambipolar to straight diffusion as the carrier concentration decreases. Disregarding this variation of drift velocity,

$$j\cdot A=\frac{n_0q\cdot\bar{v}\cdot e^{-t/\tau}}{4}A=\frac{i_{g0}e^{-t/\tau}}{4}=\frac{fi_0e^{-t/\tau}}{4}, \qquad (3)$$

where i_{g0} is the grid current, i_0 the tube current at the start of decay, f a constant factor, and τ the ion-decay time constant. Ions can be lost, as has been shown, both by diffusion and by straight transit, therefore

$$\tau=\frac{\text{Area}}{D}+\frac{\text{Av. travel length}}{\bar{c}}$$

$$\left[\bar{c}\text{ the av. velocity being proportional to }\left(\frac{T}{m}\right)^{\frac{1}{2}}\right].$$

Equating (2) and (3),

$$\tfrac{1}{4}i_0e^{-t_1/\tau}=\text{const}_1\times f_1(m)$$

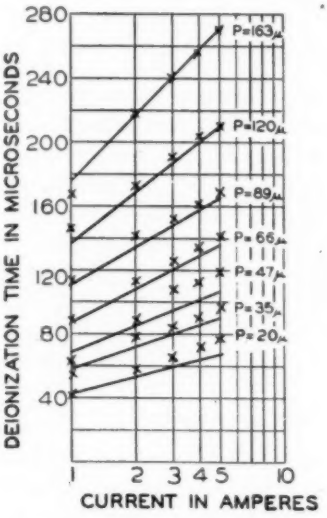


FIG. 9. Deionization time *vs* tube current in xenon. ×'s—measured points. Lines—calculated.

and

$$\text{introducing } D=\frac{\text{const}_2}{p\times f_2(m)} \text{ the recovery time}$$

$$t=\tau(\ln_e i_0-\ln_e 4\,\text{const}_1\times f_1(m))=\left[\text{area}\cdot p\cdot\text{const}_3 f_2(m)+\frac{\text{av. length}}{c}\right][\ln i_0-\ln 4\cdot\text{const}_1\times f_1(m)]$$

which for Xe gas alone leads to Eq. (1).

The information obtained on arc drops and recovery time can be useful in the design of low voltage, short deionization time tubes. It can be also of value as an indication of some factors leading to particularly short deionization times. An attempt has been made to show that a simplified concept of the processes involved, omitting secondary processes such as cumulative ionization and formation of metastable atoms is quite satisfactory for a semiquantitative understanding of the energy balance and recovery phenomena in thyratrons. The introduction of diffusion independent transit times in plasma disintegration may have a bearing upon the design of ultra-short deionization time thyratrons and upon the understanding of ion decay phenomena in the pressure region of "molecular flow."

Acknowledgment is due Mr. R. Snyder and Mr. R. Haubrich for their help with the spectrographic tests.

[8] Von W. Koch, Z. Tech. Physik **17**, 446 (1936).

[9] H. Alex Romanowitz, "Measurements, Analysis and Statistical Nature of Deionization Time in a Mercury Vapor Thyratron," Engineering Experiment Station, Bulletin No. 6, University of Kentucky, College of Engineering, December, 1947.

[10] L. Malter and E. O. Johnson, RCA Rev. **11**, 165–177 and 178–189 (1950).

APPENDIX

Observed lines of elements. (*i*—line of ionized atom, *d*—dim, + present, − absent.)

Wavelength	Tube No. 1, Fig. 4, 64μ pressure								Tube No. 2, Fig. 5, 100μ pressure							
Percent Ar	16	44	62.7	75.8	82.6	90.8	96.	100	0	18	39.4	58.8	68.9	78.8	90.3	100
Ar 3554.3	−	−	−	−	−	−	+	+	−	−	−	−	−	−	−	+
Ar 3567.7	−	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−
Ar 3606.5	−	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−
Ar 3678.3	−	−	−	−	−	−	+	+	−	−	−	−	−	−	−	−
Ar 3834.7	−	−	−	−	−	−	+	+	−	−	−	−	−	−	−	+
Ar 3949.0	−	−	−	−	+	+	+	+	−	−	−	−	+	+	+	+
Xe 3951.0	+	+	+	+	+	+	+	−	+	+	+	+	+	+	+	−
Xe 3967.6	+	+	+	+	+	+	+	−	+	+	+	+	+	+	+	−
Ar 4044.4	−	−	−	−	−	+	+	+	−	−	−	−	−	−	+	+
Xe 4078.8	+	+	+	+	+	+	+	−	+*d*	+*d*	+	+	+	+	+	−
Xe 4109.7	+	+	+	+	+	+	+	−	−	−	+	+	+	+	+	−
Xe 4116.1	+	+	+	+	+	+	+	−	+*d*	+*d*	+	+	+	+	+	−
Ar 4158.6	−	−	−	+	+	+	+	+	−	−	−	−	+	+	+	+
Ar 4164.2	−	−	−	−	−	+	+	+	−	−	−	−	−	−	+	+
Ar 4181.9	−	−	−	−	−	+	+	+	−	−	−	−	−	−	+	+
Ar 4190.7	−	−	−	−	+	+	+	+	−	−	−	−	−	−	−	−
Ar 4191.0	−	−	−	−	+	+	+	+	−	−	−	−	−	+	+	+
Xe 4193.5	+	+	+	+	+	+	+	−	+	+	+	+	+	+	+	−
Ar 4198.3	−	−	−	−	+	+	+	+	−	−	−	−	+	+	+	+
Ar 4200.7	−	−	−	+	+	+	+	+	−	−	−	−	+	+	+	+
Ar 4251.2	−	−	−	−	−	+	+	+	−	−	−	−	−	−	−	+
Ar 4259.4	−	−	−	−	+	+	+	+	−	−	−	−	+	+	+	+
Ar 4266.3	−	−	−	−	+	+	+	+	−	−	−	−	−	−	+	+
Ar 4272.2	−	−	−	−	+	+	+	+	−	−	−	−	−	+	+	+
Ar 4300.1	−	−	−	−	+	+	+	+	−	−	−	−	−	+	+	+
Ar 4333.6	−	−	−	−	+	+	+	+	−	−	−	−	−	+	+	+
Ar 4335.3	−	−	−	−	−	+	+	+	−	−	−	−	−	+	+	+
Ar 4345.2	−	−	−	−	−	+	+	+	−	−	−	−	−	−	+	+
Xe 4501.0	+	+	+	+	+	+	+	−	+	+	+	+	+	+	+	−
Ar 4510.7	−	−	−	−	−	+	+	+	−	−	−	−	−	−	+	+
Ar 4522.3	−	−	−	−	−	−	+	+	−	−	−	−	−	−	−	+
Xe 4524.7	+	+	+	+	+	+	+	−	+	+	+	+	+	+	+	−
Xe 4582.7	+	+	+	+	+	+	+	−	+*d*	+*d*	+	+	+	+	+	−
Ar 4596.1	−	−	−	−	−	−	+	+	−	−	−	−	−	−	−	+
Xe 4624.3	+	+	+	+	+	+	+	−	+	+	+	+	+	+	+	−
Ar 4628.5	−	−	−	−	−	−	+	+	−	−	−	−	−	−	−	−
Xe 4671.2	+	+	+	+	+	+	+	−	+	+	+	+	+	+	+	−
Xe 4691.0	+	+	+	+	+	+	−	−	+	+	+	+	+	+	+	−
Xe 4697.0	+	+	+	+	+	+	+	−	+	+	+	+	+	+	+	−
Ar 4702.3	−	−	−	−	−	−	+	+	−	−	−	−	−	−	−	+
Xe 4734.1	+	+	+	+	+	+	+	−	+	+	+	+	+	+	+	−
Xe 4807.0	+	+	+	+	+	+	+	−	+	+	+	+	+	+	+	−
Xe 4829.7	+	+	+	+	+	+	+	−	+	+	+	+	+	+	+	−
Xe 4843.3	+	+	+	+	+	+	+	−	+	+	+	+	+	+	+	−
Xe 4916.5	+	+	+	+	+	+	+	−	+	+	+	+	+	+	+	−
Xe 4923.1	+	+	+	+	+	+	+	−	+	+	+	+	+	+	+	−
Ar 5495.9	−	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−
Ar 5558.8	−	−	−	−	−	−	+	+	−	−	−	−	−	−	−	+
Ar 5607.0	−	−	−	−	−	−	+	+	−	−	−	−	−	−	−	+
Ar 5650.8	−	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−
Xe 5823.9	+	+	+	+	+	+	+	−	−	+*d*	+	+	+	+	+	−
Xe 5875.0	−	−	−	−	−	−	−	−	−	−	−	−	−	+	+	−
Ar 5888.0	−	−	−	−	−	+	+	+	−	−	−	−	−	−	−	−
Ar 5912.1	−	−	−	−	−	−	+	+	−	−	−	−	−	−	−	+
Xe 5931.2	−	−	−	−	−	−	−	−	−	−	−	−	−	+	+	−
Xe 5934.2	−	−	−	−	−	−	−	−	−	−	−	−	+	+	+	−
Ar 6032.1	−	−	−	−	−	+	+	+	−	−	−	−	−	−	+	+
Ar 6043.1	−	−	−	−	−	+	+	+	−	−	−	−	−	−	−	+
Ar 6059.4	−	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−
Xe 6178.3	−	−	−	−	−	−	−	−	−	−	−	+	+	+	+	−
Xe 6182.4	+*d*	+*d*	+*d*	+	+	+	+	−	−	−	+	+	+	+	+	−
Xe 6318.1	−	−	+*d*	+*d*	+	+	+	−	−	+	+	+	+	+	+	−
Ar 6416.3	−	−	−	−	−	+	+	+	−	−	−	−	−	−	−	+

Percent Ne	14	27	45	62	77.2	85.3	92.2	95.6	36	60	77.5	88.0	94
Ne 3417.9	−	−	−	−	−	−	+	+	−	−	−	−	−
Ne 3447.7	−	−	−	−	−	−	+	+	−	−	−	−	+
Ne 3454.2	−	−	−	−	−	−	+	+	−	−	−	−	+
Ne 3466.6	−	−	−	−	−	−	+	+	−	−	−	−	−
Ne 3472.6	−	−	−	−	−	−	+	+	−	−	−	−	+
Ne 3498.1	−	−	−	−	−	−	+	+	−	−	−	−	−

Appendix.—*Continued.*

Wavelength / Percent Ne	Tube No. 1, Fig. 4, 64μ pressure								Tube No. 2, Fig. 5, 100μ pressure				
	14	27	45	62	77.2	85.3	92.2	95.6	36	60	77.5	88.0	94
Ne 3501.2	−	−	−	−	−	−	+	+	−	−	−	−	−
Ne 3515.2	−	−	−	−	−	−	+	+	−	−	−	−	−
Ne 3520.5	−	−	−	−	−	+	+	+	−	−	−	−	+
Ne 3593.5	−	−	−	−	−	−	+	+	−	−	−	−	+
Xe 3951.0	+	+	+	+	+	+	+	+	+	+	+	+	+
Xe 3967.6	+	+	+	+	+	+	+	+	+	+	+	+	+
Xe 4078.8	+	+	+	+	+	+	+	+	+	+	+	+	+
Xe 4109.7	+	−	+	+	+	+	+	+	−	−	+	+	+
Xe 4116.1	+	−	+	+	+	+	+	+	−	−	+	+	+
Xe 4180.0 *i*	−	−	−	−	−	+	+	+	−	−	−	−	+
Xe 4193.5	+	+	+	+	+	+	+	+	+	+	+	+	+
? 4213	−	−	−	−	+	+	+	+	−	−	−	−	+
Xe 4245.4 *i*	−	−	−	−	+	+	+	+	−	−	−	−	+
Xe 4330.5 *i*	−	−	−	−	+	+	+	+	−	−	−	−	+
Xe 4383.2	+	−	+	+	+	+	+	+	−	−	−	−	+
Xe 4393.2 *i*	−	−	−	−	−	−	+	+	−	−	−	−	−
Xe 4395.1 *i*	−	−	−	−	−	−	+	+	−	−	−	−	−
Xe 4414.8 *i*	−	−	−	−	−	−	+	+	−	−	−	−	+
Xe 4448.1 *i*	−	−	−	−	−	−	+	+	−	−	−	−	+
Xe 4462.2 *i*	−	−	−	−	+	+	+	+	−	−	−	−	+
Xe 4501.0	+	+	+	+	+	+	+	+	+	+	+	+	+
Xe 4524.7	+	+	+	+	+	+	+	+	+	+	+	+	+
Ne 4537.8	−	−	−	−	−	−	+	+	−	−	−	−	+
Xe 4582.7	+	+	+	+	+	+	+	+	+	+	+	+	+
Xe 4603.0 *i*	−	−	−	+	+	+	+	+	−	−	−	−	+
Xe 4624.3	+	+	+	+	+	+	+	+	+	+	+	+	+
Xe 4671.2	+	+	+	+	+	+	+	+	+	+	+	+	+
Xe 4691.0	+	+	+	+	+	+	+	+	+	+	+	+	+
Xe 4697.0	+	+	+	+	+	+	+	+	+	+	+	+	+
Xe 4734.1	+	+	+	+	+	+	+	+	+	+	+	+	+
Xe 4807.0	+	+	+	+	+	+	+	+	+	+	+	+	+
Xe 4829.7	+	+	+	+	+	+	+	+	+	+	+	+	+
Xe 4843.3	+	+	+	+	+	+	+	+	+	+	+	+	+
Xe 4916.5	+	+	+	+	+	+	+	+	+	+	+	+	+
Xe 4923.1	+	+	+	+	+	+	+	+	+	+	+	+	+
Ne 5037.8	−	−	−	−	−	−	−	−	−	−	−	+	+
Xe 5292.2 *i*	−	−	−	−	+	+	+	+	−	−	−	−	+
Xe 5341.1	−	−	−	−	−	−	+	+	−	−	−	−	−
Xe 5400.5	−	−	−	−	−	−	+	+	−	−	−	−	−
Xe 5419.2 *i*	−	−	−	−	+	+	+	+	−	−	−	−	+
Ne 5689.8	−	−	−	−	−	−	−	−	−	−	−	+	+
Ne 5719.2	−	−	−	−	−	−	−	−	−	−	−	+	+
Ne 5764.4	−	−	−	−	−	−	+	+	−	−	−	−	+
Ne 5820.2	−	−	−	−	−	−	+	+	−	−	−	−	+
Xe 5823.9	+	+	+	+	+	+	+	+	+*d*	+	+	+	+
Ne 5852.5	−	−	−	−	+	+	+	+	−	−	−	−	+
Ne 5872.8	−	−	−	+	+	+	+	+	−	−	−	+	+
Ne 5881.9	−	−	−	−	+	+	+	+	−	−	−	−	+
Xe 5931.2	−	−	−	−	−	−	−	−	−	−	−	+	+
Xe 5934.2	−	−	−	−	−	−	−	−	−	−	−	+	+
Ne 5944.8	−	−	−	−	+	+	+	+	−	−	−	−	+
Ne 5974.6	−	−	−	−	−	+	+	+	−	−	−	−	+
Ne 6030.0	−	−	−	−	−	+	+	+	−	−	−	−	+
Ne 6074.3	−	−	−	−	+	+	+	+	−	−	−	−	+
Ne 6096.2	−	−	−	−	+	+	+	+	−	−	−	−	+
Ne 6143.1	−	−	−	−	+	+	+	+	−	−	−	+	+
Ne 6163.6	−	−	−	−	+	+	+	+	−	−	−	+	+
Xe 6168.3	+	−	+	+	+	+	+	+	−	−	−	+	+
Xe 6182.4	+	−	+	+	+	+	+	+	−	+	+	+	+
Xe 6198.3	−	−	−	−	+	+	+	+	−	−	−	+	+
Ne 6217.3	−	−	−	−	−	+	+	+	−	−	−	−	+
Ne 6266.5	−	−	−	−	+	+	+	+	−	−	−	−	+
Ne 6304.8	−	−	−	−	+	+	+	+	−	−	−	−	+
Xe 6318.1	+	+	+	+	+	+	+	+	+	+	+	+	+
Ne 6334.4	−	−	−	−	+	+	+	+	−	−	−	−	+
Ne 6382.0	−	−	−	−	+	+	+	+	−	−	−	−	+
Ne 6402.2	−	−	−	−	+	+	+	+	−	−	−	+	+

Percent He	10	33	42.5	62.2	85.3	93.7	33.5	47	64.0	73.5	86.0	94.0	97.0
He 3819.6	−	−	−	−	+	+	−	−	−	−	−	+	+
He 3888.6	−	−	−	+	+	+	−	−	−	+	+	+	+
Xe 3951.0	+	+	+	+	+	+	+	+	+	+	+	+	+
He 3964.7	−	−	−	−	+	+	−	−	−	−	−	+	+

Appendix.—*Continued.*

Wavelength Percent He	Tube No. 1, Fig. 4, 64μ pressure						Tube No. 2, Fig. 5, 100μ pressure						
	10	33	42.5	62.2	85.3	93.7	33.5	47	64.0	73.5	86.0	94.0	97.0
Xe 3967.6	+	+	+	+	+	+	+	+	+	+	+	+	+
He 4026.2	−	−	−	−	+	+	−	−	−	−	+	+	+
Xe 4078.8	+	+	+	+	+	+	+	+	+	+	+	+	+
Xe 4109.7	+	+	+	+	+	+	+	+	+	+	+	+	+
Xe 4116.1	+	+	+	+	+	+	+	+	+	+	+	+	+
He 4120.8	−	−	−	−	+	+	−	−	−	−	−	+	+
He 4143.7	−	−	−	−	−	+	−	−	−	−	−	−	−
Xe 4180.0 *i*	−	−	−	−	+	+	−	−	−	−	−	+	+
Xe 4193.5	+	+	+	+	+	+	+	+	+	+	+	+	+
Xe? 4214.0 *ii*	−	−	−	−	+	+	−	−	−	−	−	+	+
Xe 4238.2 *i*	−	−	−	−	+	+	−	−	−	−	−	+	+
Xe 4245.4 *i*	−	−	−	−	+	+	−	−	−	−	−	+	+
Xe 4330.5 *i*	−	−	−	−	+	+	−	−	−	−	−	+	+
Xe 4383.9	+	+	+	+	+	+	−	−	−	+	+	+	+
He 4387.9	−	−	−	−	+	+	−	−	−	−	−	+	+
Xe 4393.2 *i*	−	−	−	−	+	+	−	−	−	−	−	−	+
Xe 4395.7 *i*	−	−	−	−	+	+	−	−	−	−	−	−	+
Xe 4414.8 *i*	−	−	−	−	+	+	−	−	−	−	−	+	+
Xe 4448.1 *i*	−	−	−	−	+	+	−	−	−	−	−	+	+
Xe 4462.2 *i*	−	−	−	−	+	+	−	−	−	−	+	+	+
He 4471.5	−	−	−	−	+	+	−	−	−	−	+	+	+
Xe 4501.0	+	+	+	+	+	+	+	+	+	+	+	+	+
Xe 4524.7	+	+	+	+	+	+	+	+	+	+	+	+	+
Xe 4582.7	+	+	+	+	+	+	+	+	+	+	+	+	+
Xe 4603.0 *i*	−	−	−	+	+	+	−	−	−	+	+	+	+
Xe 4624.3	+	+	+	+	+	+	+	+	+	+	+	+	+
Xe 4671.2	+	+	+	+	+	+	+	+	+	+	+	+	+
Xe 4691.0	+	+	+	+	+	+	+	+	+	+	+	+	+
Xe 4697.0	+	+	+	+	+	+	+	+	+	+	+	+	+
He 4713.1	−	−	−	−	+	+	−	−	−	−	−	+	+
Xe 4734.1	+	+	+	+	+	+	+	+	+	+	+	+	+
Xe 4807.0	+	+	+	+	+	+	+	+	+	+	+	+	+
Xe 4829.7	+	+	+	+	+	+	+	+	+	+	+	+	+
Xe 4843.3	+	+	+	+	+	+	+	+	+	+	+	+	+
Xe 4916.5	+	+	+	+	+	+	+	+	+	+	+	+	+
He 4921.9	−	−	−	−	+	+	−	−	−	−	−	+	+
Xe 4923.1	+	+	+	+	+	+	+	+	+	+	+	+	+
He 5015.7	−	−	−	−	+	+	−	−	−	−	−	+	+
Xe 5292.2 *i*	−	−	−	+	+	+	−	−	−	−	−	+	+
Xe 5419.2 *i*	−	−	−	+	+	+	−	−	−	+	+	+	+
Xe 5823.9	+	+	+	+	+	+	+	+	+	+	+	+	+
He 5875.6	−	−	−	+	+	+	−	−	−	−	+	+	+
Xe 5931.2	−	−	−	+	+	+	−	−	−	−	+	+	+
Xe 5934.2	−	−	−	+	+	+	−	−	−	−	+	+	+
Xe 6178.3	+	+	+	+	+	+	−	+	+	+	+	+	+
Xe 6182.4	+	+	+	+	+	+	−	+	+	+	+	+	+
Xe 6318.1	+	+	+	+	+	+	−	+	+	+	+	+	+

Pore Distribution in Porous Media

H. I. MEYER
Research Department, United Gas Corporation, Shreveport, Louisiana
(Received August 7, 1952)

An attempt is made to determine the actual distribution of the pore sizes in a porous rock from the cumulative distribution of mercury injected into the rock with increasing pressure against capillary forces. The aim is to correct for those pores which could fill at a given pressure but which do not, because they are connected to the mercury source only by smaller pores. The presented correction is based on probability concepts.

INTRODUCTION

INASMUCH as there have yet to be developed unique relations between porosity and the permeability and other properties of reservoir rock, it is generally believed that at least one other parameter is needed to fully characterize porous media. One property of porous rock which it seems to be possible to obtain without undue difficulty is the size distribution of the pores. Furthermore, it seems intuitively sound to hope that the permeability, capillary pressure, and other important properties of reservoir rock might be functionally related to this distribution. It is not the purpose of this paper to consider such relationships but only to suggest a revision in the method now used to determine pore distribution. A general discussion of this phase of core analysis will be found in a recent article by Pollard and Reichertz.[1]

Ritter and Drake[2,3] proposed a method of determining pore distribution by interpreting data obtained from measuring the amount of mercury injected into a core as pressure is increased. The basis of this method is the concept that mercury is forced into smaller and smaller evacuated pores against capillary forces as the mercury pressure is increased. That such information does not directly give the pore distribution is readily seen. There may, for example, be large pores which one would expect to fill at low pressure, which have no connection with the mercury source except through smaller pores. The effect of these "ink-bottle" pores is to assign too small a portion of the pore space to the large pores and too large a part to the small pores, if the mercury injection data is taken at its face value. In this paper a method is presented for using probability theory to correct this data in an attempt to find the true pore distribution. Certain simplifying assumptions are made in order to obtain a practicable solution. Although these limitations are not stressed in the body of the paper, they are discussed in some detail in the closing paragraph.

The limitations of the mercury injection method have been recognized by others, e.g., Ritter and Drake themselves mention the effect of ink bottle pores. For a qualitative description of a reservoir rock, such a correction as is proposed here is probably not necessary; however, if pore distribution data are to be used to determine other properties of reservoir rock, an accurate pore distribution must be known. The problem of determining pertinent parameters to describe these distributions is not considered in this paper.*,[4–6] The only other work on a statistical interpretation of pore distributions that I am acquainted with is a paper by Childs and Collis-George.[7]

ANALYTICAL TREATMENT

We define the radius of a pore r to be equal to the radius of a circular tube which will just allow mercury to enter it against capillary force at the same pressure as that at which the pore will fill if connected to a mercury source.

Let $\vartheta(r)$ be equal to the fraction of the total void space occupied by pores whose radii lie in the range r to $r+dr$. The number of pores with radii in the range r to $r+dr$, which we designate $a(r)dr$, will be given by the equation

$$a dr = f\vartheta dr/Kr^3,$$

where Kr^3 is the volume of a single pore of radius r, f is the porosity of the rock.

It is necessary to designate one representative point in the interior of each pore which we shall term the "center" of that pore. For the purposes of this analysis, any convenient point will do so long as it is not too close to the wall of the pore and is consistently chosen, e.g., the center of gravity of the void space will serve. In order to develop a probability distribution, consider a volume dv so small that the probability of the centers of two or more r pores (pores whose radii lie in the radius range r to $r+dr$), being located in this volume, may be neglected. Let the probability of mr pores

[1] T. A. Pollard, and P. P. Reichertz, Bull. Am. Assoc. Petroleum Geol. **36**, 230 (1952).
[2] H. L. Ritter, and L. C. Drake, Ind. Eng. Chem. **17**, 782 (1945).
[3] H. L. Ritter and L. C. Drake, Ind. Eng. Chem. **17**, 787 (1945).

* Work of this type has been done to relate particle size distribution. See references 4–6.
[4] W. F. Cloud, Oil Weekly **103**, 26 (1941).
[5] J. C. Griffiths, Bull Am. Assoc. Petroleum Geol. **36**, 205 (1952).
[6] W. C. Krumbein and G. D. Monk, Trans. Am. Inst. Mining Met. Engrs. Petroleum Div. Reprint **146**, 151, 359–369 (1942–1943).
[7] E. C. Childs, and N. Collis-George, Proc. Roy. Soc. (London) **A201**, 392 (1950).

being located in any volume v, be $\psi_m(v, r)dr$, then

$$\psi_m(v+dv, r)dr=\psi_m(v, r)dr[1-(adr)dv] +\psi_{m-1}(v, r)dr(adr)dv,$$

if the pores are randomly distributed. Rearranging,

$$d(\psi_m dr)/dv=(adr)[(\psi_{m-1}dr)-(\psi_m dr)].$$

This equation is satisfied by Poisson's distribution, i.e.,

$$\psi_m(v, r)dr=\frac{[(adr)v]^m}{m!}\exp[-(adr)v]. \qquad (1)$$

Now we define a volume $v(r_0, r)$ such that, if an r pore is located in $v(r_0, r)$, it may be expected to be connected to an r_0 pore. This is readily done, if we assume that two pores are connected, if the center of the smaller lies in a volume adjacent to the larger pore, which is the difference between the volume of a pore whose radius equals the sum of the radii of the two pores minus the volume of the larger pore. If $r_0>r$, this volume $v(r_0, r)$ is given by the relation

$$v(r_0, r)=K(r+r_0)^3-Kr_0^3=K(3r^2r_0+3rr_0^2+r_0^3).$$

The probability of mr pores all being joined to the same r_0 pore is given by Eq. (1) with v set equal to $v(r_0, r)$. The probable number of such groups of pores which would exist in a unit volume should be given by the product of this probability times the number of $v(r_0, r)$ volumes in a unit cube (which is equal to the number of r_0 pores in a unit volume). This number, therefore, must equal

$$a(r_0)dr_0\cdot\psi_m(v(r_0, r), r)dr \quad \text{or} \quad a(r_0)dr_0\psi_m(r_0, r)dr.$$

Of particular interest is the function $\psi_0(r_0, r)dr$, which is the probability of an r_0 pore not being connected to an r pore.

Suppose a porous rock is subjected to mercury pressure such that all pores of radius $\geq r$ would fill if they had access to the mercury. Let $\Gamma(r)$ be the measured fraction of pore volume that is filled with injected mercury at this pressure, and let $\Theta(r)$ be the actual fraction of the pore volume devoted to pores of radius $\geq r$. We make the simplifying assumption that the only pores of radius $\geq r$ which will fail to fill are those that are not connected to any other pore of equal or larger radius. That is, the probability of an r_0 pore failing to fill is given in accordance with the "multiplication law of probability"[8] by the expression

$$\int_r^\infty\times\psi_0(r_0, r)dr,$$

where the symbol $\int\times$ denotes product integration.

The total number of pores of radius $\geq r_0$ which will fail to fill is given by the expression

$$\Psi(r_0, r)=\int_{r_0}^\infty a(r_0)\left\{\int_r^\infty\times\psi_0(r_0, r)dr\right\}dr_0,$$

in which the first integral sign denotes Riemann integration.

The relationship between $\Gamma(r)$ and $\Theta(r)$ is then given by the expression

$$\Gamma(r)=\Theta(r)\left[1-\Psi(r, r)\Big/\int_r^\infty a(r)dr\right]. \qquad (2)$$

This is the equation we shall use to approximate the true pore distribution $\Theta(r)$ from the distribution of injected mercury $\Gamma(r)$.

NUMERICAL APPROXIMATION

Equation (2) is not convenient for use in numerical computation. We, therefore, replace it by an approximate equation involving step functions. The radius r is considered to take on only a finite number of values $r_n<r_{n-1}<\cdots r_1$. Also set $r_{i-1}-r_i=\Delta r_i$ if $i>1$ and set $r_0=\infty$. We make the following definitions:

$$a_i=\int_{r_i}^{r_{i-1}}a(r)dr, \qquad i=1, 2\cdots$$

$$\gamma_i=\Gamma(r_i)-\Gamma(r_{i-1}), \qquad i=1, 2\cdots n$$

$$\vartheta_i=\Theta(r_i)-\Theta(r_{i-1}), \qquad i=1, 2\cdots n$$

and set up the approximations

$$\Psi(r_i, r_j)=\sum_{K=1}^{i}a_K\prod_{L=1}^{j}\psi(r_K, r_L)\Delta r_L$$

$$=\sum_{K=1}^{i}a_K\prod_{L=1}^{j}\psi_{KL},$$

and

$$\psi_{ij}=\exp[-a_iv(r_i, r_j)].$$

Substitution of these expressions into Eq. (2) gives the approximate equation

$$\Gamma(r_i)=\sum_{K=1}^{i}\gamma_K=\sum_{K=1}^{i}\vartheta_K \times\left[1-\sum_{K=1}^{i}a_K\prod_{L=1}^{i}\psi_{KL}\Big/\sum_{K=1}^{i}a_K\right]. \qquad (3)$$

Some of the quantities appearing in this expression are available *a priori*. If the pressures at which we will observe the volume of injected mercury are predetermined, and if the contact angle of mercury on the core matrix and the surface energy of the air-mercury (or vacuum-mercury) interface are known, then the r_i and v_{ij} values may be tabulated. For example, if we set the surface tension σ to equal 480 dynes/cm and the contact angle ϕ to equal 140°, then the relation

$$r_i=2\sigma\cos\phi/p(r_i)$$

[8] N. Arley, and K. Buch, *Probability and Statistics* (John Wiley and Sons, Inc., New York, 1950), paragraph 2-6.

gives the following values:

i	1	2	3	4	5	6
p_i (in psi)	25	50	100	200	500	1000
r_i (microns)	4.27	2.13	1.07	0.533	0.213	0.107
r_i^3 (microns)	77.9	9.66	1.23	0.151	0.00966	0.00123

From these we may set up the matrix

$$v_{ij} = K \begin{bmatrix} 545 & \cdots & \cdots & \cdots & \cdots & \cdots \\ 184 & 67.6 & \cdots & \cdots & \text{(symmetric)} & \cdots \\ 74.1 & 23.1 & 8.61 & \cdots & \cdots & \cdots \\ 19.4 & 9.14 & 2.86 & 1.06 & \cdots & \cdots \\ 12.0 & 3.10 & 0.87 & 0.264 & 0.0676 & \cdots \\ 6.1 & 1.0 & 0.41 & 0.111 & 0.0231 & 0.00861 \end{bmatrix}$$

cubic microns.

An iterative process based on Eq. (3) to solve for ϑ_i values from known γ_i values may be used. The steps involved are

(1) Assume a value for ϑ_i, and calculate a_i and ψ_{ij} values not already known.

(2) Using these values for ψ_{ij} and a_i, solve Eq. (3) for ϑ_i.

(3) If the computed ϑ_i value is not close enough to the assumed value, make another assumption and repeat the process.

The actual computation of ϑ_i from the known γ_i is not as formidable as Eq. (3) might make it appear. Such a process may be carried out by semi-skilled labor or may be put in a form to be handled by a computing machine.

SAMPLE PROBLEM

In order to demonstrate the effect of this correction, a specific example was worked out. The original data were published by Burdine *et al.*[9] The core in question was designated by them as Sample No. 826A. The mercury injection data are plotted on their published Fig. 2. This core is natural sandstone, cylindrical of length 2.74 cm and diameter 2.37 cm. Its laboratory permeability was measured to be 440 millidarcies and its porosity 0.237.

The significant data, both measured and computed, are presented in Table I.

The data were divided into six groups corresponding to six injection pressures 25, 50, 100, 200, 500, and 1000 psi. Actual computation time was of the order of two hours. In this case, the probabilities of the smallest pores not being connected became vanishingly small, and so the injection at higher pressures does not need correction. At low pressures there is considerable difference between γ_i and ϑ_i.

TABLE I.

i	Fraction of volume filled w/mercury (from Burdine)	Injected differential fraction —	Corrected differential fraction (Eq. 3)
1	0.03	0.03	0.14
2	0.20	0.17	0.18
3	0.46	0.26	0.17
4	0.58	0.12	0.10
5	0.67	0.09	0.08
6	0.73	0.06	0.06

[9] Burdine, Gournay, and Reichertz, J. Petroleum Technol. **2**, 195 (1950).

DISCUSSION AND CONCLUSION

It might be well to list the various assumptions, either announced or implicit, in the preceding discussion. The most important are

(1) The assumed randomness of pore distribution.

(2) The assumption that the volume of a pore may be expressed as Kr^3, where K is a constant.

(3) The ignoring of the fact that pores are mutually exclusive in deriving Poisson's distribution.

(4) The assumption that the center of an r pore lying in $v(r_0, r)$ will assure its being connected to an r_0 pore.

(5) The assumption that only the pores which are isolated from all equal or larger pores will fail to fill.

This is a formidable list, and some defense had better be made for each simplification. Statement (1) possibly excludes from this treatment small samples of such rocks as vuggy limestone, while homogeneous sandstones and oolitic limestones would be acceptable. The critical point in (2) is not that K exists but that it is a constant, an assumption that is almost certainly not true. However, K as a constant appears only under the integral sign and cancels out. Actually then the assumption is not that K is a constant but that its mean value in one integral is equal to that in the other.

Inasmuch as we only used the "zero" term in Poisson's distribution, it is hoped that assumption (3) does not introduce serious error. The criterion for connectedness in (4) is at least valid for spherical pores and seems intuitively sound for others. The last assumption (5) is probably the most difficult to justify; actually it implies that we are not taking into consideration isolated groups of mutually connected pores. One defense of this simplification is the idea that such groups may be considered to be one large pore, which concept has the effect of reducing the numerical value of K. This is sort of taking advantage of K, however, and it is better to admit that, in the interest of simplicity, we are ignoring a rather difficult subject, that of aggregates in a random distribution.[10] In any event we do correct for single isolated pores and hope that this is the principal correction required.

As mentioned in the introduction, the purpose of trying to find the true pore distribution in a reservoir rock is the hope that we may be able to use physical properties of porous rocks to determine its character as an oil and gas producer. Such an approach to core analysis is possibly methodologically sounder than the present methods which largely consist of the conducting of small-scale flow experiments under controlled laboratory conditions.

[10] C. Mack, Proc. Cambridge Phil. Soc. **46**, 285 (1950).

Evaporation of Germanium Films from a Carbon Crucible*

K. LEHOVEC,† J. ROSEN, A. MACDONALD, AND J. BRODER
Signal Corps Engineering Laboratories, Fort Monmouth, New Jersey
(Received September 2, 1952)

Equipment for evaporation of germanium in vacuum at various measurable rates is described. The activation energy of evaporation is 3.5 ev, which is in close agreement with recent data by Searcy.

INTRODUCTION

As a preliminary to the task of studying the properties of germanium films evaporated onto a substrate of single crystal germanium or silicon, a series of experiments was undertaken to determine the rate of germanium evaporation in vacuum as a function of temperature (approximately 1100°C to 1550°C).

DISCUSSION

Germanium was deposited in vacuum (~0.1-μ pressure) on a mica sheet[1] from a resistance heated carbon boat. The boat was shaped by cuttings so that the central portion of the boat containing the germanium was at a high and constant temperature. The temperature was measured by a platinum *versus* platinum–rhodium thermocouple, placed in the bottom wall of the boat. The fanned-out ends of the carbon boat were copper plated and held in copper sleeves that were bolted to water-cooled copper electrodes (see Fig. 1).

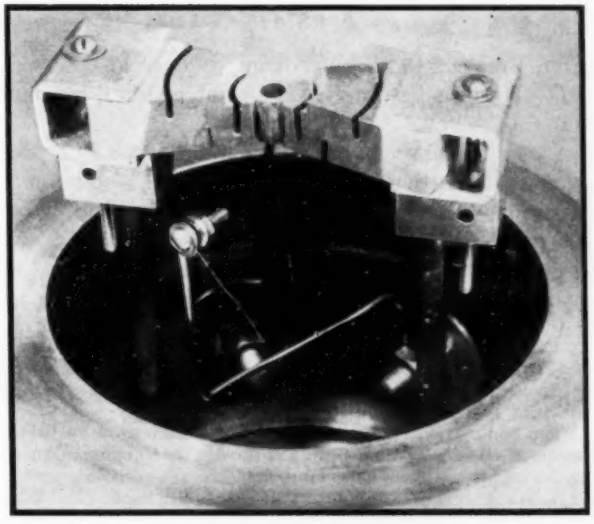


FIG. 1. Carbon crucible used in evaporation of germanium.

The transmission through the mica sheet, as detected by a germanium photocell, was recorded as a function of time. Each transmission curve consists of a series of maxima and minima. Figure 2 shows a typical curve of relative transmission *versus* time. The rate of deposition p_0 was determined from the period τ between two successive maxima by the equation

$$p_0=\lambda/2\eta\tau.\ddagger$$

For λ the value 1.7 microns was used. This value corresponds to the spectral absorption edge of germanium. The index of refraction for germanium at $\lambda=1.7$ microns is $\eta=4.6$.[2]

At temperatures above 1450°C, the rate of deposition of germanium becomes so fast that the transmission

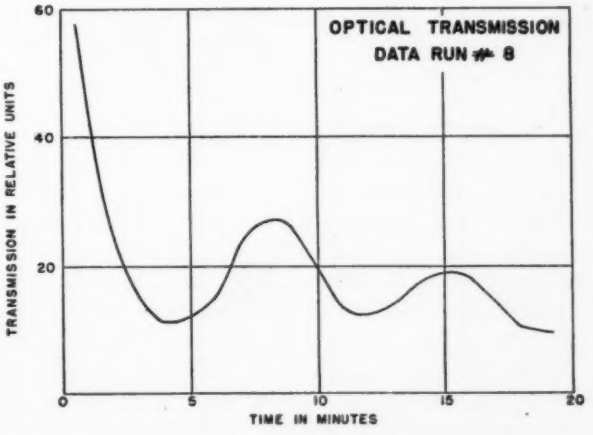


FIG. 2. Sample curve of relative transmission *versus* time.

maxima and minima followed each other too rapidly to be recorded by a human observer. Since a suitable automatic recorder was not available, the deposited germanium film was weighed. From these data together with the total time of evaporation, a mass deposition rate p_m (given in g/sec, see Table I) was calculated. The mass deposition rate is a value averaged across the whole mica sheet. The optical method previously described gives a deposition rate only for the center of the mica sheet just above the carbon crucible. The calibration factor necessary to convert average deposition

* Presented at the Semiconductor Conference sponsored by the Institute of Radio Engineers-American Institute of Electrical Engineers in Urbana, Illinois, June 1952.

† Now at Sprague Electric Company, North Adams, Massachusetts.

[1] The mica sheet was located 34.5 mm above the center portion of the carbon crucible, the liquid germanium resting about 3 mm below the top of the crucible opening.

‡ The subscript 0 indicates that this deposition rate is derived from optical transmission data to differentiate it from p_m, which is the deposition rate derived from a measurement of the weight of the germanium film deposited.

[2] W. H. Brattain and H. B. Briggs, Phys. Rev. 75, 1705 (1949).

TABLE I. Data on the evaporation rates of germanium. (Distance from liquid germanium to substrate 37.5 mm).

Run number	Temperature °C	Weight of film deposited[b] (mg)	Rate of evaporation p_m from mass data (g/sec)	Rate of evaporation p_0 from optical data (Å/sec)	Period τ[c] (sec)	Ratio p_0/p_m (cm/g)
4	1114	3.1	0.252×10^{-6}	0.27	6750	1.1×10^{-2}
3	1187	8.0	1.99	1.6	1125	0.81
8	1259	5.1	3.40	4.1	450	1.2
12	1259	5.6	3.46	3.6	510	1.0
1	1259	18.8	9.49	6.6	281	0.69
11	1259	7.2	10.0	12	150	1.2
2	1332	11.1	12.3	13	146	1.1
5	1404	10.8	40.0	46	40	1.1
6	1477	10.5	175.0	180[a]		
7	1550	5.9	246	250[a]		
9	1550	9.2	307	310[a]		
10	1550	4.8	320	320[a]		

[a] These data are discussed in the text and are marked on the curve by surrounding them with squares.
[b] The deposited film decreases in thickness from center to outside. In order to get film thickness in center multiply value in this column by calibration factor 0.01 cm/g.
[c] τ = time in seconds between maxima in optical data.

rate to the deposition rate applicable at the center of the mica sheet was obtained from a comparison of optical data and weight data at the lower temperatures,

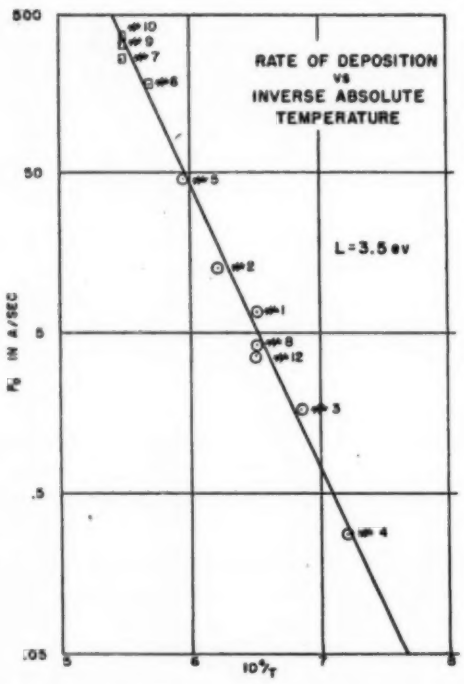


FIG. 3. Evaporation rate curve of germanium (rate in angstroms per second *versus* inverse temperature in degrees Kelvin). Rate plotted on logarithmic scale.

where both were available. This calibration factor, the average of the ratio p_0/p_m at the lower temperatures, was used in calculating the rate of deposition in the center of the mica sheet at the higher temperatures, where only weight data were available. An attempt was made to determine the absolute value of the rate of deposition directly above the crucible by weighing the germanium deposited in a small central area. It was assumed that the deposit was homogeneous in this central portion.

The rate of evaporation was found to be little dependent on the size of the germanium melt used. This was probably due to the fact that the area of evaporation was limited by the boat opening and the system geometry.

With the exception of runs 8 and 11 in which abnormally small and large charges, respectively, were used, germanium charges of about 50 milligrams were chosen.

Occasional observation on evaporation at higher residual pressures than 0.1 micron make it appear that the rate of evaporation does not depend critically on the vacuum, within the temperature and pressure ranges investigated here.

From the rate of evaporation as a function of temperature (Fig. 3), one obtains an activation energy of 3.5 ev. The vapor pressure *vs* temperature relationship for germanium has been tabulated by Brewer.[3] These vapor pressure data were apparently derived from the boiling point temperature, $T_B=2980°K$[4]; the use of the rule of Pictet and Trouton, $S=L/T_B=23$ calories/degree Kelvin; and the equation

$$\log_{10}P=L/4.574T+S/4.574.$$

From Brewer's data, by plotting a curve $\log_{10}P$ *versus* $1/T$ (all points lie exactly on a straight line), one finds that $L=2.92$ ev.

A plot of Searcy's[5] deposition rate *versus* $1/T$ gives an activation energy $L=3.3$ ev, a similar plot for his vapor pressure *versus* $1/T$ has a slope corresponding to a value for L of 3.4 ev. Our activation energy of 3.5 ev is in excellent agreement with Searcy's results, and apparently Trouton's rule cannot be applied to derive the boiling point temperature of germanium.

[3] L. Brewer, "Chemistry and metallurgy of miscellaneous materials," Paper 3, p. 31, L. E. Quill, Editor, *Thermodynamics* (McGraw-Hill Book Company, Inc., 1950).
[4] The boiling point temperature of germanium used by Brewer was obtained from a value quoted in the *Metals Handbook* (American Society for Metals, Cleveland, 1939).
[5] A. W. Searcy, J. Am. Chem. Soc. **74**, 4789 (1952). (Published after our work was concluded.)

JOURNAL OF APPLIED PHYSICS VOLUME 24, NUMBER 5 MAY, 1953

Propagation of Acoustic Waves in a Liquid-Filled Cylindrical Hole, Surrounded by an Elastic Solid

EDWARD V. SOMERS *
Magnolia Petroleum Company, Field Research Laboratories, Dallas, Texas
(Received September 24, 1952)

The theory of elastic wave propagation in systems where discontinuities in the physical properties occur along concentric cylindrical surfaces is developed for an impulsive ring-shaped source. This source is mathematically similar to the point source discussed by Lamb and Pekeris. The field problem is solved for a liquid-filled column, surrounded by an elastic solid, representing a basic model applicable to well-logging problems.

I. INTRODUCTION

MANY theoretical papers have been written on the subject of wave propagation in systems composed of elastic solids and liquids. Lamb[1] in 1904 treated the propagation of waves over the surface of an isotropic elastic-solid half-space, the wave source being an arbitrary application of vertical force at a point in the solid. Pekeris[2] in 1948 extended Lamb's treatment to wave propagation in two and three liquid layers, with an arbitrary wave source located in the first layer. Pekeris' theory adequately explained the shape and variation of amplitude of a received pressure pulse having its origin in an exploding charge in a shallow water-covered area, e. g., continental-shelf area of the ocean. Press and Ewing[3–8] have worked out numerous extensions of the Lamb-Pekeris treatment. In particular, one of their contributions has been an analysis of the propagation of waves in a liquid layer overlying an elastic solid layer, useful in analyzing submarine disturbances.

In the above analyses, the discontinuities in physical properties occur along planes parallel to one another; this includes the point source, which has a discontinuity in one direction only. In cases where the discontinuities occur along concentric circular cylindrical surfaces, the theoretical treatment has not been completed to the same satisfactory degree. Existing treatments, many of which use approximating simplifications, are confined to determination of the phase-velocity curves arising from the boundary conditions imposed at the solid-liquid interface for harmonically oscillating sources. These treatments compare satisfactorily with results obtained experimentally.

Lamb[9] in 1898, extending the work of Korteweg,[10] treated the effect of the elastic yielding of the tube walls on the propagation of sound in a liquid-filled tube. Two cases were investigated by Lamb: (1) the case of a thin-walled tube; (2) the case of an infinitely thick-walled tube. Green[11] in 1923 obtained a result identical with Lamb's, for the thin-walled cylinder. Gronwall[12] in 1927 extended Lamb's original treatment to tubes not completely filled with liquid; he also extended Lamb's results by not restricting his series expansions to a single term, as Lamb had done. Field[13] in 1931 expanded the problem to high frequencies, in order to explain phase velocity to acoustic velocity ratios greater than one, which were noted experimentally at these frequencies by Boyle and Froman;[14] he found such a solution, apparently overlooked originally by Korteweg.[10] Jacobi[15] in 1948 treated five simple models theoretically. He included data that indicated that two of these simple models were each in agreement with phase velociteis experimentally measured in a physical system approximating the assumptions of the model. Because of his Fourier expansions, Jacobi finds an infinite number of modes rather than the two discussed by Field. He discusses the problem of determining the field produced by an arbitrary excitation, in terms of characteristic functions originating from the Sturm-Liouville system for the radial coordinate. As indicated by Gronwall's work, the characteristic functions must be determined with the axial coordinate rather than the radial. Further, to correctly satisfy the boundary conditions at the solid-liquid surface, one needs a fixed correspondence between the axial boundary conditions in the solid and the liquid. The problem of establishing a source for determining the arbitrary coefficients and then of extending the harmonic solution to that of an arbitrary excitation is a further complication.

The present paper establishes a complete solution to the field problem for a particular model. Essentially, the

* Present address: Westinghouse Electric Corporation, Research Laboratories, East Pittsburgh, Pennsylvania.

[1] H. Lamb, Trans. Roy. Soc. (London) A203, 1–42 (1904).
[2] Ewing, Worzel, and Pekeris, Geol. Soc. Am. Mem. 27 (1948).
[3] F. Press and M. Ewing, Trans. Am Geophys. Union 27, 163 (1948).
[4] F. Press and M. Ewing, Geophysics, 13, 404 (1948).
[5] Press, Ewing, and Tolstoy, Bull. Seis. Soc. Am. 40, 111 (1950).
[6] F. Press and M. Ewing, Geophysics 15, 426 (1950).
[7] F. Press and M. Ewing, J. Appl. Phys. 22, 892 (1951).
[8] F. Press and M. Ewing, Trans. Am. Geophys. Union 32, 673 (1951).
[9] H. Lamb, Manchester Mem. 42, 3 (1898).
[10] D. J. Korteweg, Ann. Physik, 5, 525 (1878); quoted by Lamb, reference 9.
[11] H. G. Green, Phil. Mag. 45, 907 (1923).
[12] T. H. Gronwall, Phys. Rev. 30, 71 (1927).
[13] G. S. Field, Can. J. Research 5, 131 (1931).
[14] R. W. Boyle and D. Froman, Nature 126, 602 (1930).
[15] W. J. Jacobi, J. Acoust. Soc. Am. 21, 120 (1949).

model extends Gronwall's treatment to a solid infinite along the axis of the liquid column, with a source of the type discussed by Lamb and Pekeris established within the liquid column. The source used here is a ring-shaped line source instead of the point source used in the plane-layered cases. The solution for longitudinal and shear potentials is obtained in integral form for the source oscillating harmonically. The evaluation of the integrals then follows the work of Pekeris and can be extended to arbitrary excitation of the source.

II. THE ACOUSTIC MODEL

Consider the propagation of elastic waves in a system consisting of a circular cylindrical column of liquid, surrounded by an elastic solid. The density of the liquid is given by ρ_1 and its incompressibility by λ_1; the solid is characterized by a density ρ_2, an incompressibility factor λ_2, and a shear modulus μ_2. The corresponding three wave velocities are given by

$$\begin{aligned} \alpha_1&=(\lambda_1/\rho_1)^{\frac{1}{2}},\\ \alpha_2&=\left(\frac{\lambda_2+2\mu_2}{\rho_2}\right)^{\frac{1}{2}},\\ \beta_2&=\left(\frac{\mu_2}{\rho_2}\right)^{\frac{1}{2}}. \end{aligned} \tag{1}$$

For simple harmonic motion ($\epsilon^{i\omega t}$) the wave equations in the liquid (subscript 1) and the solid (subscript 2) are given by

$$\begin{aligned} \left(\nabla^2+\frac{\omega^2}{\alpha_1^2}\right)\varphi_1&=0,\\ \left(\nabla^2+\frac{\omega^2}{\alpha_2^2}\right)\varphi_2&=0,\\ \left(\nabla^2+\frac{\omega^2}{\beta_2^2}\right)\psi_2&=0, \end{aligned} \tag{2}$$

where

$$\nabla^2=\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}+\frac{\partial^2}{\partial z^2}. \tag{3}$$

The longitudinal and shear potentials, per Lamb's[1] procedure, are defined in terms of the component displacements l_r and l_z:

$$\begin{aligned} (l_r)_1&=\partial\varphi_1/\partial r,\\ (l_z)_1&=\partial\varphi_1/\partial z,\\ (l_r)_2&=(\partial\varphi_2/\partial r)+(\partial^2\psi_2/\partial r\partial z),\\ (l_z)_2&=(\partial\varphi_2/\partial z)+(\partial^2\psi_2/\partial z^2)+(\omega^2/\beta_2^2)\psi_2. \end{aligned} \tag{4}$$

The normal and tangential stresses are given by

$$\begin{aligned} N_r&=\lambda\nabla\cdot\mathbf{l}+2\mu\frac{\partial l_r}{\partial r},\\ N_z&=\lambda\nabla\cdot\mathbf{l}+2\mu\frac{\partial l_z}{\partial z},\\ T_\varphi&=\mu\left(\frac{\partial l_r}{\partial z}+\frac{\partial l_z}{\partial r}\right). \end{aligned} \tag{5}$$

If we select the origin of the coordinate system in the center of the liquid column and place a wave source symmetrical with respect to the plane $z=0$, we need only consider the half-space, $z>0$. The following boundary conditions apply:

(a) The vertical displacement along the plane, $z=0$, is zero, and the normal stress at z infinite is zero, in both the liquid and the solid;

$$\left.\begin{aligned} (l_z)_1&=0\\ (l_z)_2&=0 \end{aligned}\right\},\quad (z=0) \tag{6}$$

$$(N_z)_1=(N_z)_2=0,\quad (z\text{ infinite}).$$

(b) The normal stress across the solid-liquid interface is continuous, as is the normal displacement. The tangential stress is zero;

$$\left.\begin{aligned} (N_r)_1&=(N_r)_2\\ (l_r)_1&=(l_r)_2\\ (T_\varphi)_2&=0 \end{aligned}\right\},\quad (r=a) \tag{7}$$

where

$$\begin{aligned} N_r&=\lambda\nabla^2\varphi+2\mu\frac{\partial l_r}{\partial r}\\ T_\varphi&=\mu\left(2\frac{\partial^2\varphi_2}{\partial r\partial z}+2\frac{\partial^3\psi_2}{\partial r\partial z^2}+\frac{\omega^2}{\beta_2^2}\frac{\partial\psi_2}{\partial r}\right). \end{aligned} \tag{8}$$

To establish a source for the problem, select the cylindrical surface, $r=b$, within the liquid, so that the liquid space is divided into an annular region, with a prime denoting the potential, and a circular cylindrical region with the potential is unprimed. Along this cylindrical surface, continuity of pressure is required,

$$\nabla^2\varphi_1'=\nabla^2\varphi_1. \tag{9}$$

For a section of this cylindrical surface, $z<|z_0|$, let there be established a radially directed source, such that there is a constant-valued discontinuity in the radial displacement. This is given by

$$\frac{\partial\varphi_1'}{\partial r}-\frac{\partial\varphi_1}{\partial r}=\begin{cases}(2\pi bz_0)^{-1}, & (z<z_0)\\ (4\pi bz_0)^{-1}, & (z=z_0)\\ 0 & (z>z_0).\end{cases} \tag{10}$$

As z_0 approaches zero, the discontinuity in the radial displacement approaches that of a line source in such a manner that its integral over the cylindrical surface, $r=b$, from $z=0$ to z infinite, is unity,

$$\int_0^{2\pi}\int_0^{\infty}\left(\frac{\partial\varphi_1'}{\partial r}-\frac{\partial\varphi_1}{\partial r}\right)bd\theta dz=2\pi b\int_0^{z_0}\frac{1}{2\pi bz_0}dz=1. \quad (11)$$

Equation (10) can be rewritten in terms of a Fourier-cosine integral as follows:

$$\frac{\partial\varphi_1'}{\partial r}-\frac{\partial\varphi_1}{\partial r}=\frac{2}{\pi}\int_0^{\infty}\frac{\sin qz_0}{4bqz_0}\cos qzdq. \quad (12)$$

Equations (2), (6), (7), (9), and (12) constitute the mathematical problem to be solved. Equation (12) is expressed in a form agreeing with that of the superposed solution of the potential function arising from the boundary condition, Eq. (6), a Fourier-cosine integral form. Contrasted to this solution, Gronwall's solutions for models of finite extent in the direction, are Fourier-cosine or sine expansions.

III. THE SOLUTION OF THE MODEL

Complete solutions to Eqs. (2) obtained by separating variables and superposing solutions, satisfying the boundary conditions given by Eq. (6), follow:

$$\begin{aligned}
\varphi_1&=\frac{2\epsilon^{i\omega t}}{\pi}\int_0^{\infty}A_1(r)\cos qzdq,\\
\psi_2&=\frac{2\epsilon^{i\omega t}}{\pi}\int_0^{\infty}B_1(r)\sin qzdq,\\
\varphi_1'&=\frac{2\epsilon^{i\omega t}}{\pi}\int_0^{\infty}C_1(r)\cos qzdq,\\
\varphi_1&=\frac{2\epsilon^{i\omega t}}{\pi}\int_0^{\infty}D_1(r)\cos qzdq,
\end{aligned} \quad (13)$$

where

$$\begin{aligned}
A_1(r)&=EK_0(\eta r), \quad (r\geq b),\\
B_1(r)&=DK_0(\delta r), \quad (r\geq b),\\
C_1(r)&=BI_0(\xi r)+CK_0(\xi r), \quad (b\leq r\leq a),\\
D_1(r)&=AI_0(\xi r), \quad (0\leq r\leq b).
\end{aligned} \quad (14)$$

The quantities ξ, η, δ are defined by

$$\begin{aligned}
\rho^2&=q^2-\frac{\omega^2}{\alpha_1^2},\\
\eta^2&=q^2-\frac{\omega^2}{\alpha_2^2},\\
\delta^2&=q^2-\frac{\omega^2}{\beta_2^2}.
\end{aligned} \quad (15)$$

In terms of the arbitrary factors in Eq. (13), we can re-express Eq. (12) as follows:

$$\frac{\partial\{C_1(r)\}}{\partial r}-\frac{\partial\{D_1(r)\}}{\partial r}=G=\frac{\sin qz_0}{4bqz_0}. \quad (16)$$

Applying the boundary conditions given by Eqs. (7) and (9) under the integral signs of Eq. (25), and combining with Eq. (16) the values of A, B, C, D, and E given in Eqs. (14) can be determined from the five resulting equations. Thus,

$$\begin{aligned}
\varphi_1&=\frac{2\epsilon^{i\omega t}}{\pi}\int_0^{\infty}\frac{\Delta_A}{\Delta}I_0(\xi r)\cos qzdq,\\
\psi_2&=\frac{2\epsilon^{i\omega t}}{\pi}\int_0^{\infty}\frac{\Delta_B}{\Delta}K_0(\delta r)\sin qzdq,
\end{aligned} \quad (17)$$

where, for example, Δ is the determinant of the coefficients of A, B, etc., and Δ_A is the determinant arising by replacement of the A coefficients by the nonhomogeneous parts of the five equations.

Since it is expected that most experimental measurements are made along the axis of the liquid column, let us confine our solution to φ_1. Solutions for the other three potential functions are obtained similarly. For φ_1,

$$\begin{aligned}
\Delta_A=G\Big\{&\lambda_1(\xi^2-q^2)^2[I_0(\xi b)K_0(\xi a)-K_0(\xi b)I_0(\xi a)]\\
&\cdot\left[\frac{\omega^2}{\beta_2^2}\eta\delta K_1(\eta a)K_1(\delta a)\right]\\
&+\xi(\xi^2-q^2)[K_0(\xi b)I_1(\xi a)-I_0(\xi b)K_1(\xi a)]\\
&\cdot\left[\left\{\frac{\omega^2}{\beta_2^2}-2q^2\right\}\{(\lambda_2+2\mu_2)(\eta^2-q^2)\right.\\
&+2\mu_2q^2\}\left\{\frac{K_0(\eta a)}{\eta K_1(\eta a)}\right\}\\
&\left.-4\mu_2q^2\delta^2\frac{K_0(\delta a)}{\delta K_1(\delta a)}-\frac{4\mu_2q^2}{a}\right][\eta\delta K_1(\eta a)K_1(\delta a)]\Big\};
\end{aligned} \quad (18)$$

and for φ_1, φ_1', φ_2, and ψ_2,

$$\begin{aligned}
\Delta=\Big\{&\left[\frac{\omega^2}{\beta_2^2}-2q^2\right][(\lambda_2+2\mu_2)(\eta^2-q^2)+2\mu_2q^2]\\
&\times\left[\xi I_1(\xi a)\frac{K_0(\eta a)}{\eta K_1(\eta a)}\right]+\left[4q^2\mu_2\delta^2\frac{K_0(\delta a)}{\delta K_1(\delta a)}\xi I_1(\xi a)\right]\\
&+\left[\frac{\omega^2}{\beta_2^2}2\mu_2\xi\frac{I_1(\xi a)}{a}\right]+\left[\frac{\omega^2}{\beta_2^2}\lambda_1(\xi^2-q^2)I_0(\xi a)\right]\Big\}\\
&\cdot\{\eta\delta K_1(\eta a)K_1(\delta a)\}\cdot\left\{\frac{\xi^2-q^2}{b}\right\}.
\end{aligned} \quad (19)$$

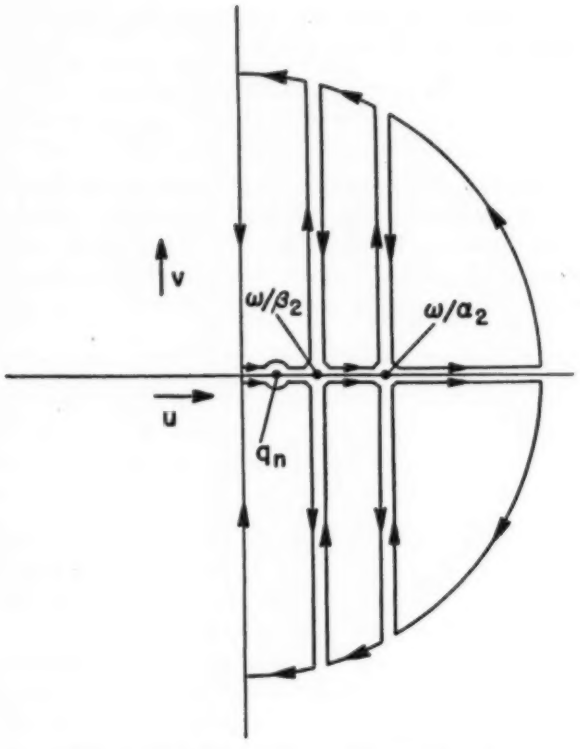


FIG. 1. Path of integration in the complex q plane.

In event the frequency ω of impressed wave is small, ξ is small, and accurate approximations can be made for the following functions in Δ and Δ_A:

$$I_0(\xi a)\sim 1.0; \quad I_1(\xi a)\sim \xi a/2. \tag{20}$$

The value of Δ is then reduced to that given by Lamb[9] as Eq. (61) of his 1898 paper. For very low frequencies, the phase velocity to acoustic velocity ratio (α/α_1) can be quickly obtained from Eq. (19); it is another of Lamb's results,

$$\left(\frac{\alpha}{\alpha_1}\right)^2=\frac{\mu_2}{\lambda_1+2\mu_2}, \tag{21}$$

and is the basis of many undeveloped schemes for well-logging the shear modulus of rock formation.

IV. EVALUATION OF THE INTEGRAL SOLUTION

Integrals of the type given in Eq. (17) have been investigated by Lamb,[1] Pekeris,[2] and others. The integral is considered as the sum of two integrals,

$$\varphi_1=\frac{\epsilon^{i\omega t}}{\pi}\int_0^\infty \frac{\Delta_A}{\Delta}I_0(\xi r)\epsilon^{iqz}dq + \frac{\epsilon^{i\omega t}}{\pi}\int_0^\infty \frac{\Delta_A}{\Delta}I_0(\xi r)\epsilon^{-iqz}dq \tag{22}$$

$$=\frac{\epsilon^{i\omega t}}{\pi}\int_0^\infty \frac{M(q)}{\Delta(q)}[\epsilon^{iqz}+\epsilon^{-iqz}]dq.$$

The procedure is to transform the path of integration into the complex q plane, $q=u+\epsilon v$, so that the solution φ_1 is then contained within the sum of two contour integrals,

$$0=\int_C \epsilon^{iqz}\frac{M(q)}{\Delta(q)}dq+\int_{C'}\epsilon^{-iqz}\frac{M(q)}{\Delta(q)}dq. \tag{23}$$

[From Eqs. (23) through (39), we suppress the $\epsilon^{i\omega t}$ factor and use the absolute value only.] Assuming the roots of $\Delta(q)=0$ to lie on the positive real axis, the integrals will have nonzero contributions for all roots, $\{q_n\}$, and for the branch points $q=\omega/\alpha_2$ and $q=\omega/\beta_2$. The C and C' contours are taken in the first and fourth quadrants, as shown in Fig. 1.

The integral contribution to φ_1 are now, considering C only,

$$(\varphi_1)_C=i\int_\infty^0 \epsilon^{-vz}\frac{M(iv)}{\Delta(iv)}dv+\int_0^\infty \epsilon^{iuz}\frac{M(u)}{\Delta(u)}du - p_1-p_2-\pi i\sum_{n=1}^{\infty}(\text{Residues of } q_n), \tag{24}$$

where p_1 and p_2 are the branch-point contributions along the barrier $u=\omega/\beta_2$ and $u=\omega/\alpha_2$, respectively.

Similarly for C',

$$(\varphi_1)_{C'}=i\int_{-\infty}^0 \epsilon^{vz}\frac{M(iv)}{\Delta(iv)}dv+\int_0^\infty \epsilon^{-iuz}\frac{M(u)}{\Delta(u)}du - p_1'-p_2'-\pi i\sum_{n=1}^{\infty}(\text{Residues of } q_n). \tag{25}$$

Since $M(iv)=M(-iv)$ and $\Delta(iv)=\Delta(-iv)$, we find, by adding $(\varphi_1)_C$ to $(\varphi_1)_{C'}$, that

$$\int_0^\infty(\epsilon^{iuz}+\epsilon^{-iuz})\frac{M(u)}{\Delta(u)}du=p_1+p_1'+p_2+p_2' + \pi i\sum_{n=1}^{\infty}(\text{Residue } q_n)+(\text{Residue } q_n)'. \tag{26}$$

The evaluation of p_1 can be taken for large values of z, so that, with $q=\omega/\beta_2+iv$, only small values of v need be considered. Accordingly, the values of δ, given on the left-hand side and right-hand side of the barrier, respectively, are

$$\delta_{\text{L.H.S.}}=\left(\frac{2v\omega}{\beta_2}\right)^{\frac{1}{2}}\epsilon^{-3\pi i/4}=\delta_1, \qquad \delta_{\text{R.H.S.}}=\left(\frac{2v\omega}{\beta_2}\right)^{\frac{1}{2}}\epsilon^{i\pi/4}=-\delta_1. \tag{27}$$

The value of η, the same on both sides of the ω/β_2 barrier, is given by

$$\eta_1=(\omega^2/\beta_1^2-\omega^2/\alpha_1^2)^{\frac{1}{2}}. \tag{28}$$

The value of p_1 is then approximated by

$$p_1 \sim i\int_0^\infty \left[\frac{\Delta_A(\delta_1,\eta)}{\Delta(\delta_1,\eta)} - \frac{\Delta_A(-\delta_1,\eta)}{\Delta(-\delta_1,\eta)}\right] I_0(\xi r)\epsilon^{-vz}dv. \quad (29)$$

The same integral is obtained for p_1' in the fourth quadrant. Two useful mathematical identities can be employed at this point:

$$\begin{aligned} K_0(-\delta_1 a) &= K_0(\delta a), \\ K_1(-\delta_1 a) &= -K_1(\delta_1 a) - i\pi I_1(\delta_1 a). \end{aligned} \quad (30)$$

Inspection of Δ_A/Δ reveals it to be of the form

$$\frac{\Delta_A(\delta)}{\Delta(\delta)} = \frac{A_1 + \delta K_1(\delta a)B_1}{A_2 + \delta K_1(\delta a)B_2}, \quad (31)$$

so that from Eqs. (30) and (31),

$$\frac{\Delta_A(-\delta)}{\Delta(-\delta)} = \frac{\Delta_A(\delta) + \delta i\pi I_1(\delta a)B_1}{\Delta(\delta) + \delta i\pi I_1(\delta a)B_2} = \frac{\Delta_A(\delta) + \delta B_1'}{\Delta(\delta) + \delta B_2'}. \quad (32)$$

By substituting Eqs. (31) and (32) into (29),

$$p_1 \sim (1-i)(\omega/\beta_2)^{\frac{1}{2}} \int_0^\infty v^{\frac{1}{2}} f(v)\epsilon^{-vz}dv, \quad (33)$$

where

$$f(v) = \frac{(B_1' - (\Delta_A/\Delta)B_2')}{\Delta + \delta_1 B_2'} I_0(\xi r). \quad (34)$$

Equation (33) for large values of z can be evaluated directly by the saddle-point method,[16]

$$\int_0^\infty v^{\frac{1}{2}} f(v)\epsilon^{-vz}dv \sim \frac{\Gamma(\frac{3}{2})}{Z^{\frac{1}{2}}} f(0). \quad (35)$$

The evaluation of p_2 and p_2' follows the same line of reasoning and gives a similar result,

$$p_2 \sim (1-i)(\omega/\alpha_2)^{\frac{1}{2}} \int_0^\infty v^{\frac{1}{2}} g(v)\epsilon^{-vz}dv, \quad (36)$$

where $g(v)$ is derived in the same way as $f(v)$, with the result in Eq. (34) having the value δ replaced by the value η. Here the values η_2 and δ_2, used along the ω/α_2 barrier, are

$$\begin{aligned} \eta_{\text{L.H.S.}} &= \left(\frac{2v\omega}{\alpha_2}\right)^{\frac{1}{2}} \epsilon^{-3\pi i/4} = \eta_2, \\ \eta_{\text{R.H.S.}} &= \left(\frac{2v\omega}{\alpha_2}\right)^{\frac{1}{2}} \epsilon^{i\pi/4} = -\eta_2, \\ \delta_2 &= \left(\frac{\omega_2^2}{\alpha_2^2} - \frac{\omega_2^2}{\beta_2^2}\right)^{\frac{1}{2}}. \end{aligned} \quad (37)$$

[16] R. Courant and D. Hilbert, *Methoden der Mathematischen Physik* (Interscience Publishers, Inc., New York, 1943).

The normal-mode evaluation for the C contour equals

$$\begin{aligned} \sum_{n=1}^\infty (\text{Residues } q_n) &= \sum_{n=1}^\infty \frac{\Delta_A(qn)}{\partial[\Delta(qn)]/\partial q} I_0(\xi_n r)\epsilon^{iq_n z} \\ &= \sum_{n=1}^\infty \Phi(q_n) I_0(\xi_n r)\epsilon^{iq_n z}. \end{aligned} \quad (38)$$

For the C' contour,

$$\sum_{n=1}^\infty (\text{Residues } q_n)' = \sum_{n=1}^\infty \Phi_1(q_n) I_0(\xi_n r)\epsilon^{-iq_n z}. \quad (39)$$

We obtain, from Eqs. (22), (26), (35), (36), (38), and (39),

$$\begin{aligned} \varphi_1 = 2^{\frac{1}{2}}\epsilon^{-i(\pi/4)} \frac{\Gamma(\frac{3}{2})}{Z^{\frac{1}{2}}} \left\{ f(0)\left(\frac{\omega}{\beta_2}\right)^{\frac{1}{2}} + g(0)\left(\frac{\omega}{\alpha_2}\right)^{\frac{1}{2}} \right\} \\ + 2\pi i \sum_{n=1}^\infty \Phi(q_n) I_0(\xi_n r)\cos q_n z. \end{aligned} \quad (40)$$

It can be seen that for large values of z the compressional and shear contributions to the wave motion from the solid are negligible compared with the water-wave contribution because of their $\frac{3}{2}$-powered decay with z. For large distances from the source, one would expect to see the normal-mode solutions only.

V. EXTENSION TO AN ARBITRARY EXCITATION OF THE SOURCE

The extension of the solution at a point far from the source, i.e., normal-mode solutions, follows directly from Pekeris[2] and Press and Ewing.[6] Their method is reiterated here, simply for completeness.

The group velocity for propagation in a dispersive medium is defined by the formula

$$U = \frac{d\omega}{dq} = \frac{d(\alpha q)}{dq} = \alpha + q\frac{d\alpha}{dq}, \quad (41)$$

where α is the phase velocity and q is the wave number.

For the medium selected, the elastic wave velocities are computed. For different values of ω, the roots of the characteristic equation (19) are computed, i.e., the curve for the values of the phase velocity α *versus* frequency is obtained numerically. From this curve the group-velocity curve can be obtained by using Eq. (41).

Following Pekeris, we represent the time variation $f(t)$ of the initial disturbance by its Fourier transform,

$$g(\omega) = \int_{-\infty}^\infty \epsilon^{-i\omega t} f(t)dt. \quad (42)$$

Then

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^\infty \epsilon^{i\omega t} g(\omega)d\omega. \quad (43)$$

The potential φ_1, referring to the steady-state potential in Eq. (39), can be written as

$$\varphi_1(r,z,t)=\sum_{n=1}^{\infty}\int_{-\infty}^{\infty}g(\omega)\epsilon^{i(\omega t-q_n z)}\Phi(q_n)I_0(\xi_n r)d\omega. \quad (44)$$
(on C')

If, again following Pekeris, the disturbance can be represented by

$$f(t)=\begin{cases}\epsilon^{-\lambda t}, & t>0\\ 0, & t<0,\end{cases} \quad (45)$$

so that

$$g(\omega)=1/\lambda+i\omega, \quad (46)$$

we find

$$\varphi_1(r,z,t)=\frac{1}{2\pi}\sum_{n=1}^{\infty}\int_{-\infty}^{\infty}\frac{\epsilon^{i(\omega t-q_n z)}}{\lambda+i\omega}\Phi(q_n)I_0(\xi_n r)d\omega. \quad (47)$$
(on C')

Rewriting Eq. (47),

$$\varphi_1(r,z,t)=\frac{1}{2\pi}\sum_{n=1}^{\infty}\int_{-\infty}^{\infty}\frac{Q_n(\omega)\epsilon^{if(\omega)}}{\lambda+i\omega}d\omega. \quad (48)$$
(on C')

For z large, the integrand of Eq. (48) is rapidly oscillatory, owing to the $\epsilon^{if(\omega)}$ factor. Consequently, the principal contribution to the integral arises at the points (ω_0) where $f'(\omega)=0$; Kelvin's method of stationary phase[17] can be used to evaluate Eq. (48),

$$\begin{aligned}f(\omega)&=\omega t-q_n z,\\ f'(\omega)&=t-z\frac{dq_n}{d\omega}=t-\frac{z}{U_n}=0,\\ f''(\omega)&=-\frac{z}{U_n^2}\frac{dU_n}{d\omega},\end{aligned} \quad (49)$$

where U_n is the group velocity of the nth mode.

By expanding $f(\omega)$ in the vicinity of ω_0; $\omega=\omega_0+y$,

$$f(\omega)=f(\omega_0)-\frac{z}{2}q_n''(\omega_0)y^2-\frac{z}{6}q_n'''(\omega_0)y^3-\cdots. \quad (50)$$

From Eqs. (48), (49), and (50), we obtain

$$\varphi_1(r,z,t)\sim\frac{1}{2\pi}\sum_{n=1}^{\infty}\frac{Q_n(\omega_0)}{\lambda+i\omega_0}\epsilon^{if(\omega_0)}\left(\frac{2\pi}{|f''(\omega_0)|}\right)^{\frac{1}{2}}\epsilon^{\pm i(\pi/4)}. \quad (51)$$
(on C')

The plus sign ($\epsilon^{+i\pi/4}$) is used for $f''(\omega_0)<0$, and the minus sign ($\epsilon^{-i\pi/4}$), for $f''(\omega_0)>0$, insuring the convergence of the integral.

[17] H. Jeffreys and B. S. Jeffreys, *Methods of Mathematical Physics* (Cambridge University Press, Cambridge, 1946), Chap. XVII.

For the steady-state potential, Eq. (38), using $\epsilon^{iq_n z}$ instead of $\epsilon^{-iq_n z}$, the evaluation centers around $-\omega_0$ instead of $+\omega_0$, so that with $Q_n(\omega_0)=Q_n(-\omega_0)$, $f''(\omega_0)=f''(-\omega_0)$,

$$\varphi_1(r,z,t)\sim\frac{1}{2\pi}\sum_{n=1}^{\infty}\frac{Q_n(\omega_0)}{\lambda-i\omega_0}\times\epsilon^{-i\omega_0 t+q_n z}\left(\frac{2\pi}{|f''(\omega_0)|}\right)^{\frac{1}{2}}\epsilon^{\pm i\pi/4}. \quad (52)$$

The combined contribution of both solutions yields

$$\varphi_1(r,z,t)=\sum_{n=1}^{\infty}\frac{Q_n(\omega_0)\cos\{\omega_0 t-zq_n(\omega_0)-\tan^{-1}(\lambda/\omega_0)\pm\pi/4\}}{[2\pi|f''(\omega_0)|(\lambda^2+\omega_0^2)]^{\frac{1}{2}}}. \quad (53)$$

The $+\pi/4$ holds if $f''(\omega_0)<0$ and $-\pi/4$ holds if $f''(\omega_0)>0$.

Per Eqs. (41) and (49), U_n, $dU_n/d\omega$, ω_0, etc., are evaluated from the phase-velocity and group-velocity curves, calculated from the characteristic equation, $\Delta(\omega)=0$, Eq. (19). These quantities then permit evaluation of $\varphi_1(r,z,t)$, as given in Eq. (53).

Where the group velocity is a minimum, $U=U_{\min}$, so that $f''(\omega)\sim 0$, it is necessary to make an extended approximation involving the second derivative of the group velocity. Here, the development of Pekeris is followed again, yielding the result

$$\varphi_1(r,z,t)=\sum_{n=1}^{\infty}\frac{Q_n(\omega_0)}{(\lambda^2+\omega_0^2)^{\frac{1}{2}}}\left[\frac{2E(s)\cos\{f(\omega_0)-\tan^{-1}(\lambda/\omega_0)\}}{3^{\frac{1}{2}}z^{\frac{1}{2}}[q_n'''(\omega_0)]^{\frac{1}{2}}}\right], \quad (54)$$

where

$$s=\frac{2\sqrt{2}z}{3(|q'''(\omega_0)|)^{\frac{1}{2}}}\left(\frac{1}{U_n}-\frac{1}{U_{n\min}}\right)^{\frac{3}{2}}, \quad (55)$$

and

$$\begin{aligned}E(s)&=s^{\frac{1}{2}}[J_{-\frac{1}{3}}(s)+J_{\frac{1}{3}}(s)], && \begin{cases}t<r/U_{\min}\\ t>r/U_{\max}\end{cases}\\ E(s)&=s^{\frac{1}{2}}[I_{-\frac{1}{3}}(s)+I_{\frac{1}{3}}(s)], && \begin{cases}t>r/U_{\min}\\ t<r/U_{\max}\end{cases}\end{aligned} \quad (56)$$

This solution, given by Eq. (54) and called the Airy phase by Pekeris, decays as the $z^{-\frac{1}{2}}$ power, compared with the $z^{-\frac{1}{3}}$ powered decay for the solution for wave trains traveling at other group velocities [given by Eq. (53)].

VI. NOTE ON PHASE VELOCITY GREATER THAN WATER VELOCITY

The above problem has been solved for phase velocities less than water velocity,

$$\alpha<\alpha_1<\beta_2<\alpha_2. \quad (57)$$

Although it is noted that solutions exist for

$$\alpha_1<\alpha<\beta_2<\alpha_2, \tag{58}$$

no attempt has been made to include these in this paper.

Since completing the writing of this paper, two recent references have come to the author's attention. Biot,[18] working with the infinitely thick-walled case, has removed the large wavelength approximation of Lamb[9] and has examined the wave propagation in great detail, including in his work plots of the phase-velocity and group-velocity *vs* frequency curves for the coupled fluid-solid system. Fay,[19] working with liquid-filled steel tubes, has similarly computed phase-velocity dispersion curves over a range of values of the ratio of the inner and outer tube radii.

ACKNOWLEDGMENT

The author expresses his appreciation to Dr. D. H. Clewell and Dr. W. J. Yost for their interest and assistance in the writing of this paper and to the Magnolia Petroleum Company for its assistance in publishing it. It was during the author's stay at Magnolia's Field Research Laboratories in Dallas that much of the thought leading to the paper was developed.

[18] M. A. Biot, J. Appl. Phys. **23**, 997 (1952).

[19] R. D. Fay, J. Acoust. Soc. Am. **24**, 459 (1952).

JOURNAL OF APPLIED PHYSICS VOLUME 24, NUMBER 5 MAY, 1953

Subharmonic Oscillations in Nonlinear Systems*

CHIHIRO HAYASHI
Department of Electrical Engineering, University of Kyoto, Kyoto, Japan
(Received September 2, 1952)

This paper deals with the subharmonic oscillations which occur in systems with nonlinear restoring force. It is first shown that the order of the subharmonics has a close connection to the form of the nonlinear characteristics. The subharmonic oscillation of order $\frac{1}{3}$, i.e., the oscillation whose fundamental frequency is one-third that of the applied force, is then particularly investigated for the cases in which the nonlinear characteristics are expressed by (1) cubic and (2) quintic functions. In both cases the stability problem of the periodic solution is discussed in detail according to the stability criterion given previously by the author. The analysis reveals that in the latter case (2) the second higher-harmonic of the subharmonic, i.e., the oscillation of order $\frac{2}{3}$ builds up with negative damping. This is verified by experiments in which the oscillation of order $\frac{2}{3}$ causes the collapse of the original subharmonic oscillation.

1. INTRODUCTION

SUBHARMONIC oscillations, whose frequency is a fraction $1/\nu$ ($\nu=2, 3, 4, \cdots$) of that of the applied force may frequently occur in nonlinear systems. We consider the differential equation

$$\frac{d^2v}{d\tau^2}+2\delta\frac{dv}{d\tau}+f(v)=B\cdot\cos\nu\tau, \tag{1.1}$$

in which 2δ is a constant damping coefficient and $f(v)$ is a term characterizing the nonlinear restoring force. It will be noticed that, since the period of the applied force is $2\pi/\nu$, the subharmonic oscillation of order $1/\nu$ has a period 2π and may be expressed by a linear combination of $\sin\tau$ and $\cos\tau$.

In the following sections, we shall first consider the relationship between the nonlinear characteristics expressed by the term $f(v)$ and the order $1/\nu$ of the subharmonic oscillations. Then, the subharmonic oscillation of order $\frac{1}{3}$ is particularly investigated with special attention directed to the stability of the periodic oscillation. (The discussion, throughout the present paper, is confined to the steady state of oscillations. A brief discussion on the transient state has previously been given by the author.[1])

2. RELATIONSHIP BETWEEN THE NONLINEAR CHARACTERISTICS AND THE ORDER OF SUBHARMONIC OSCILLATIONS

In order to investigate this we shall consider the following polynomial for the restoring force $f(v)$, i.e.,

$$f(v)=c_1v+c_2v^2+c_3v^3+\cdots, \tag{2.1}$$

where c_1, c_2, c_3, $\cdots$ are constants determined by the nonlinear characteristics and subjected, without loss of generality, to the condition

$$c_1+c_2+c_3+\cdots=1.\dagger \tag{2.2}$$

In so far as we deal with the steady state of oscillations, the periodic solution of Eq. (1.1) may be approximated by

$$v=v_0+x\sin\tau+y\cos\tau+w\cos\nu\tau, \tag{2.3}$$

* Submitted to the Faculty of Engineering of Kyoto University in partial fulfillment of the requirements for the degree of Doctor of Science, September, 1952.

[1] C. Hayashi, J. Appl. Phys. **24**, 198 (1953).

† We can choose v such that $f(v)=1$ when $v=1$, and hence Eq. (2.2) results.

in which the constant term v_0, the subharmonic oscillation $x\sin\tau+y\cos\tau$, and the oscillation having the applied frequency $w\cos\nu\tau$ are only considered on account of their prime importance.‡ Following Mandelstam and Papalexi,[2] the amplitude w may further be approximated by

$$w=\frac{1}{1-\nu^2}B. \tag{2.4}$$

This approximation is legitimate in the case when the nonlinearity is small. But, as will be shown later, the relation (2.4) is a fairly good approximation even when the departure from linearity is large.

Substituting Eq. (2.3) into Eq. (1.1), and equating to zero the coefficients of $\sin\tau$ and $\cos\tau$ separately, we have the following cases according to the forms of the nonlinear characteristics (2.1).

Case 1. When the nonlinearity is given by $f(v)=c_1v+c_3v^3$.

Under circumstances when the nonlinearity is symmetrical, i.e., $f(v)$ is odd in v, the constant term v_0 in Eq. (2.3) is usually discarded, and, for $\nu=2, 4, 5, \cdots$, the above-mentioned substitution leads to

$$\left.\begin{aligned}&[1-\tfrac{3}{4}(x^2+y^2)-\tfrac{3}{2}w^2]x+ky=0,\\&[1-\tfrac{3}{4}(x^2+y^2)-\tfrac{3}{2}w^2]y-kx=0,\end{aligned}\right\} \tag{2.5}$$

where $k=2\delta/c_3$. Multiplying the first equation by y, the second by x, and subtracting the second product from the first, we obtain

$$k(x^2+y^2)=0.$$

This means that the amplitude of the subharmonic oscillation is zero as long as the damping is present (i.e., $k\neq0$). Hence the subharmonic oscillations of orders $\frac{1}{2}, \frac{1}{4}, \frac{1}{5}, \cdots$ cannot occur in this case. However, as will be investigated in the following section, real roots of x and y which are not simultaneously zero may be obtained in the case when $\nu=3$, and some of them are maintained in stable states.§ Hence, the conclusion that the subharmonic oscillation of order $\frac{1}{3}$ may occur in the case when the nonlinear term c_3v^3 is contained in Eq. (2.1).

Case 2. When the nonlinearity is given by $f(v)=c_1v+c_2v^2+c_3v^3$.

The nonlinearity is unsymmetrical and the constant term v_0 in Eq. (2.3) must be considered. The subharmonic oscillation of order $\frac{1}{2}$ occurs predominantly in this case. It takes a considerable length to analyze the subharmonic oscillation of order $\frac{1}{2}$, so that the detailed discussion will be deferred to another paper.‖

Case 3. When the nonlinearity is given by $f(v)=c_1v+c_5v^5$.

Although the nonlinear term c_3v^3 is absent in this case, the subharmonic oscillation of order $\frac{1}{3}$ is maintained. The investigation will be given in Sec. 4. The subharmonic oscillation of order $\frac{1}{5}$ is also maintained in this case.

From the foregoing considerations, it may be deduced in general that the presence of the term $c_\nu v^\nu$ in Eq. (2.1) is the sufficient condition for the occurrence of the subharmonic oscillation of order $1/\nu$. However, this is not the necessary condition as we see in Case 3. It may also be concluded that the subharmonic oscillation of order $1/\nu$ does not occur when the highest degree of the power of the nonlinear terms in Eq. (2.1) is less than ν.

3. SUBHARMONIC OSCILLATION OF ORDER 1/3 WITH THE NONLINEAR CHARACTERISTIC $u=c_1v+c_3v^3$¶

We shall hereafter investigate the subharmonic oscillation of order $\frac{1}{3}$ in detail.

(1) Nondissipative Case

Putting

$$\left.\begin{aligned}&\nu=3, \quad k=0,\\&c_2=c_4=c_5=\cdots=0,\end{aligned}\right\}$$

in Eqs. (1.1) and (2.1), and remembering the condition (2.2), we have the following equation for the subharmonic oscillation of order $\frac{1}{3}$, i.e.,

$$\frac{d^2v}{d\tau^2}+v=c_3(v-v^3)+B\cos3\tau. \tag{3.1}$$

By the use of Eq. (2.4), the periodic solution will be given by

$$\left.\begin{aligned}&v=x\sin\tau+y\cos\tau+w\cos3\tau,\\&\text{with}\qquad w=\frac{1}{1-3^2}B=-\frac{1}{8}B.\end{aligned}\right\} \tag{3.2}$$

Substituting Eq. (3.2) into Eq. (3.1), and equating to zero the coefficients of $\sin\tau$ and $\cos\tau$ separately, we have

$$\left.\begin{aligned}&x[1-\tfrac{3}{4}(x^2+y^2)-\tfrac{3}{2}w^2]=-\frac{3w}{4}\cdot2xy,\\&y[1-\tfrac{3}{4}(x^2+y^2)-\tfrac{3}{2}w^2]=-\frac{3w}{4}(x^2-y^2).\end{aligned}\right\} \tag{3.3}$$

‡ It is tacitly assumed that the damping coefficient 2δ is not so large that the term containing $\sin\nu\tau$ may be discarded in Eq. (2.3). This approximation is usually applied without serious error; see for instance, Z. Physik **73**, 229 (1932).

[2] L. Mandelstam and N. Papalexi, Z. Physik **73**, 227 (1932).

§ These real roots represent the states of equilibrium which do not always exist actually, but are only able to last as long as they are stable.

‖ It is only mentioned here that, by a closer investigation in which the term v_0 is taken into account, the subharmonic oscillation of order $\frac{1}{2}$ may narrowly take place even in the case when $c_2=0$ in Eq. (2.1). This result is also verified by experiments, though the range of occurrence is considerably restricted. Therefore, strictly speaking, the conclusion in Case 1 should be modified in this respect.

¶ As will be mentioned in Sec. 5, this characteristic represents physically the relationship between the magnetic flux v and the magnetizing current $f(v)$ of an iron-cored reactor, and so the constants c_1 and c_3 have values between zero and unity.

Multiplying the first equation by y, the second by x, and subtracting the products so formed, we obtain

$$x=0 \quad \text{or} \quad x=\pm\sqrt{3}y.$$

For each value of x we have two pairs of roots, hence we obtain six pairs of equilibrium states in all. But we are satisfied with the investigation of the equilibrium states only for which $x=0$, because, as will be shown later, the other states have the same amplitude, but differ in phase by $2\pi/3$ or $4\pi/3$ radians in τ, *viz.*, one or two complete cycles of the applied force. This is a plausible result since we are concerned with the subharmonic oscillation of order $\frac{1}{3}$.

For $x=0$, assuming that $y\neq 0$ in the second equation of (3.3), we obtain

$$y^2+yw+2w^2-4/3=0. \tag{3.4}$$

This equation shows the relationship between the amplitude w (which is here assumed to be proportional to the applied force B) and the amplitude y of the subharmonic oscillation and is plotted as a part of an ellipse in Fig. 1.

We are now to determine the stability of the equilibrium states given by Eq. (3.4). The condition for the stability of the periodic solution is obtained by examining whether any variation from this state caused by a sufficiently small perturbation attenuates or not with the lapse of time. If this variation is denoted by ξ, then writing $v(\tau)+\xi$ for $v(\tau)$ in Eq. (3.1), we obtain the variational equation

$$\frac{d^2\xi}{d\tau^2}+(c_1+3c_3v^2)\xi=0.$$

Remembering that $x=0$ in Eq. (3.2), the equilibrium state is represented by

$$v=y\cos\tau+w\cos 3\tau.$$

Substituting this into the above variational equation, we obtain the following equation of Hill's type, namely,

$$\left.\begin{aligned}&\frac{d^2\xi}{d\tau^2}+[\theta_0+2\theta_1\cos 2\tau+2\theta_2\cos 4\tau+2\theta_3\cos 6\tau]\xi=0,\\ &\text{with}\\ &\theta_0=c_1+\tfrac{3}{2}c_3(y^2+w^2),\\ &\theta_1=\tfrac{3}{4}c_3(y^2+2yw),\\ &\theta_2=\tfrac{3}{2}c_3yw,\\ &\theta_3=\tfrac{3}{4}c_3w^2.\end{aligned}\right\} \tag{3.5}$$

Since w and y are determined by Eqs. (3.2) and (3.4), respectively, the parameters θ's in Eqs. (3.5) may readily be calculated.

According to the previous investigation,[3] the stability

[3] C. Hayashi, J. Appl. Phys. **24**, 344 (1953).

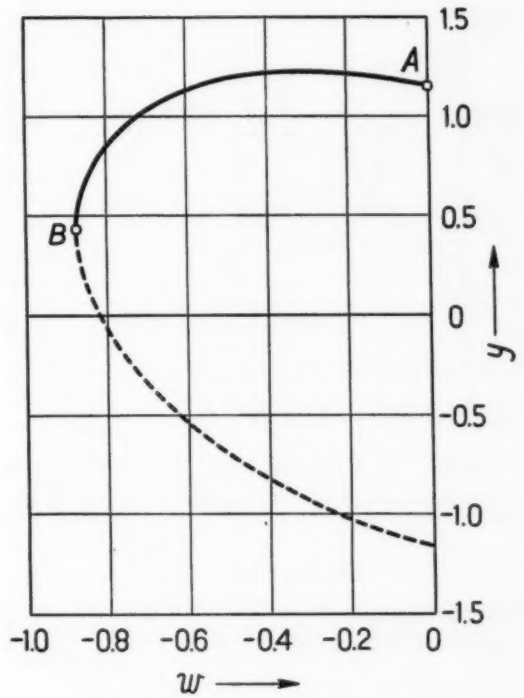


FIG. 1. Amplitude characteristic of the $\frac{1}{3}$-harmonic oscillation (nonlinearity by cubic function).

condition of the first approximation is given by

$$\left.\begin{aligned}&(\theta_0-n^2+\theta_n)(\theta_0-n^2-\theta_n)>0,\\ \text{or}\quad &|\theta_n|<|\theta_0-n^2|, \quad (n=1, 2, 3, \cdots).\end{aligned}\right\} \tag{3.6}$$

The condition for $n=1$ determines the stability against building up of the unstable oscillation having the same frequency as that of the subharmonic. By the use of Eqs. (3.4) and (3.5) this condition leads to

$$2y>-w. \tag{3.7}$$

Hence, the equilibrium states represented by the dotted-line curve in Fig. 1 are unstable and do not exist actually. Figure 2 shows the loci of $\theta_n(n=1, 2, 3)$ which are drawn by varying the value of w (or B) for the limiting case of $c_1=0$ and $c_3=1$. As expected from the stability condition for $n=1$, θ_1 enters into the first unstable region in the dotted-line interval ab. At the critical points a and b,

$$2y=-w,$$

and these points correspond respectively to A and B in Fig. 1.

It is clear from Fig. 2 that neither θ_2 nor θ_3 enters into its corresponding unstable region. Thus we see that, since the conditions (3.6) for $n=2$ and $n=3$ are satisfied, only the condition (3.7) is sufficient to determine the stability in the present case.** Hence, in the end, we

** If $\theta_m(m\geqq 2)$ enters into the mth unstable region, the stability conditions in Eq. (3.6) are not satisfied for $n=m$, and the unstable oscillation of order $m/3$ will take place. If this oscillation grows to be predominant, the original subharmonic oscillation of order 1/3 will no longer be maintained. An example for such a case ($m=2$) will be given later in Sec. 4.

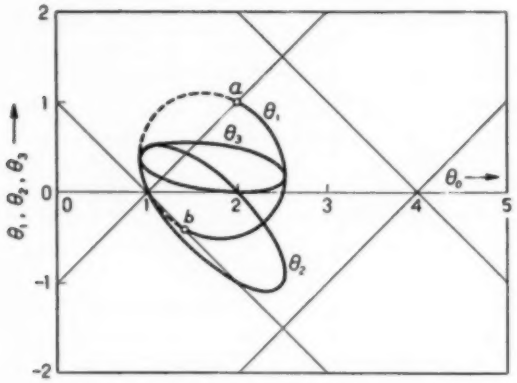


FIG. 2. Loci of θ's in Eqs. (3.5) with varying w.

may conclude that the part of the curve (in Fig. 1) which lies between A and B represents the stable oscillations.

(2) Dissipative Case

When the damping is considered in nonlinear systems, the fundamental equation takes the form

$$\frac{d^2v}{d\tau^2}+2\delta\frac{dv}{d\tau}+c_1v+c_3v^3=B\cos 3\tau, \tag{3.8}$$

and the periodic solution is expressed by††

$$\left.\begin{aligned} &v=x\sin\tau+y\cos\tau+w\cos 3\tau,\\ \text{with}\quad &w=-\frac{B}{8}. \end{aligned}\right\} \tag{3.9}$$

Substituting Eq. (3.9) into Eq. (3.8), and equating to zero the coefficients of $\sin\tau$ and $\cos\tau$ separately, we have

$$\left.\begin{aligned} &Ax+ky=-\frac{3w}{4}\cdot 2xy,\\ &Ay-kx=-\frac{3w}{4}(x^2-y^2),\\ \text{where}\quad &A=1-\tfrac{3}{4}(x^2+y^2)-\tfrac{3}{2}w^2,\\ &k=\frac{2\delta}{c_3}. \end{aligned}\right\} \tag{3.10}$$

Squaring and adding the first two equations, we obtain

$$A^2+k^2=\left(\frac{3w}{4}\right)^2r^2,\quad r^2=x^2+y^2, \tag{3.11}$$

or, by putting $r^2=R$ and $w^2=W$,

$$\frac{9}{16}R^2+\left(\frac{27}{16}W-\frac{3}{2}\right)R+\left(\frac{9}{4}W^2-3W+k^2+1\right)=0. \tag{3.12}$$

†† See same reference concerning the periodic solution of Eq. (2.3) in Sec. 2.

Hence the relationship between W and R, and consequently the amplitude r of the subharmonic oscillation is determined and depicted in Fig. 3a and 3b for various values of k.

The components x, y of the amplitude r may readily be obtained as follows. By Eqs. (3.10) and (3.11), we have

$$x^3-\frac{3}{4}Rx-\frac{kR}{3w}=0,$$

and

$$y^3-\frac{3}{4}Ry-\frac{AR}{3w}=0.$$

Hence the components x, y are given by

$$\left.\begin{aligned} &x=-r\cos\theta,\ -r\cos(\theta+120°),\ -r\cos(\theta+240°),\\ &y=-r\sin\theta,\ -r\sin(\theta+120°),\ -r\sin(\theta+240°),\\ \text{with}\quad &\cos 3\theta=-\frac{4k}{3wr}. \end{aligned}\right\} \tag{3.13}$$

The stability problem may be treated in the same manner as in the preceding case. The equation which characterizes the small variation from the periodic states of equilibrium is

$$\frac{d^2\xi}{d\tau^2}+2\delta\frac{d\xi}{d\tau}+(c_1+3c_3v^2)\xi=0.$$

By the well-known transformation $\xi=e^{-\delta\tau}\cdot\eta$, this leads to

$$\frac{d^2\eta}{d\tau^2}+[c_1-\delta^2+3c_3v^2]\eta=0.$$

Substituting the periodic solution (3.9), we obtain

$$\left.\begin{aligned} &\frac{d^2\eta}{d\tau^2}+[\theta_0+2\theta_1\cos(2\tau-\epsilon_1)+2\theta_2\cos(4\tau-\epsilon_2)\\ &\qquad\qquad+2\theta_3\cos(6\tau-\epsilon_3)]\eta=0,\\ &\text{where}\\ &\theta_0=c_1-\delta^2+\tfrac{3}{2}c_3(x^2+y^2+w^2),\\ &\theta_n^2=\theta_{ns}^2+\theta_{nc}^2,\quad \epsilon_n=\arctan\frac{\theta_{ns}}{\theta_{nc}},\quad (n=1,2,3),\\ &\theta_{1s}=\tfrac{3}{2}c_3x(y-w),\quad \theta_{1c}=\tfrac{3}{4}c_3(-x^2+y^2+2yw),\\ &\theta_{2s}=\tfrac{3}{2}c_3xw,\quad \theta_{2c}=\tfrac{3}{2}c_3yw,\\ &\theta_{3s}=0,\quad \theta_{3c}=\tfrac{3}{4}c_3w^2. \end{aligned}\right\} \tag{3.14}$$

The stability condition in dissipative systems is given by

$$(\theta_0-n^2)^2+2(\theta_0+n^2)\delta^2+\delta^4>\theta_n^2,\quad (n=1,2,3,\cdots). \tag{3.15}$$

The condition for $n=1$ is obtained by substituting θ_0

and θ_1 of Eqs. (3.14) into Eq. (3.15), and further by virtue of Eqs. (3.10) and (3.12), we ultimately find

$$R+\frac{3}{2}W-\frac{4}{3}>0. \tag{3.16}$$

Hence the equilibrium states represented by the dotted-line curves in Fig. 3 are unstable and do not actually exist. Figure 4 shows the loci of $\theta_n(n=1,2,3)$ which are drawn by varying the value of w (or B) for the case of $c_1=0$, $c_3=1$, and $\delta=0.1$. As illustrated in the figure, the boundary curve of the first unstable region is given by a hyperbola,‡‡ and θ_1 enters into this region in the dotted-line interval ab. At the critical points a and b, we have

$$R+\frac{3}{2}W-\frac{4}{3}=0,$$

and these points correspond, respectively, to A and B in Fig. 3. It is also obvious from Fig. 4 that neither θ_2 nor θ_3 enters into its corresponding unstable region. Hence, in the end, the condition (3.16) for $n=1$ is sufficient to determine the stability of equilibrium states in the case when the nonlinearity is characterized by a cubic function.

The following are some remarks on the approximation in the foregoing analysis.

As shown in Fig. 1 or 3, the stable range of equilibrium states is interposed between the critical points A and B at which w has its limiting values. Since w is proportional to B by Eqs. (3.2) or (3.9), these stability limits are obtained when the applied force also has its limiting values. This is a plausible result from the physical point of view.

However, as mentioned before, the approximation in Eq. (2.4) is legitimate as long as the nonlinearity is

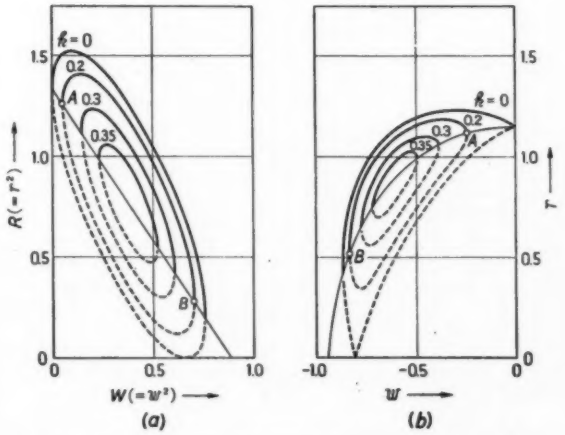


FIG. 3. a. Relationship between W and R in Eq. (3.12). b. Amplitude characteristic of the ½-harmonic oscillation.

‡‡ In the case of $\delta=0$, this hyperbola is reduced to the straight lines (drawn by chain lines in Fig. 4) which intersect the abscissa at the point (1, 0) [see Fig. 2].

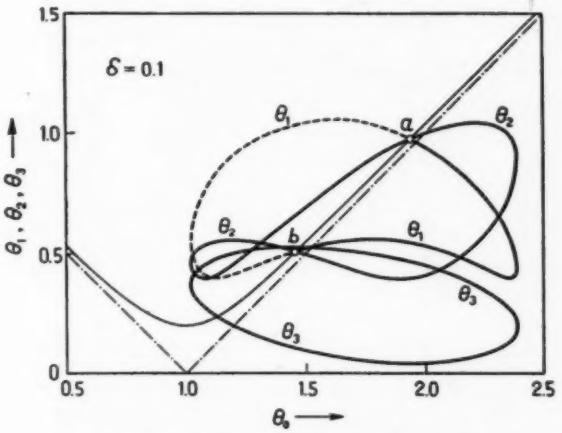


FIG. 4. Loci of θ's in Eqs. (3.14) with varying w.

small, *viz.*, $c_3 \ll c_1$ in the foregoing analysis. Thus, it might seem improper to apply this assumption in the case when $c_1=0$, $c_3=1$ in Figs. 2 or 4. But it will be explained below that the relation of Eq. (2.4) may be applied with a fairly good approximation even when the nonlinearity is predominant, and also that the stability limits mentioned above are given by the condition that the applied force B has its limiting values.

For the sake of simplicity, we consider the nondissipative case. Putting $c_1=0$, $c_3=1$ in Eq. (3.1), we have

$$\frac{d^2v}{d\tau^2}+v^3=B\cos 3\tau.$$

Substituting the periodic solution

$$v=y\cos\tau+w\cos 3\tau,$$

and equating to zero the coefficients of $\cos\tau$ and $\cos 3\tau$ separately, we obtain

$$\left.\begin{aligned} y^2+yw+2w^2-\frac{4}{3}&=0,\\ \frac{1}{4}y^3+\frac{3}{2}y^2w+\frac{3}{4}w^3-9w&=B. \end{aligned}\right\} \tag{3.17}$$

Figure 5 is obtained by plotting Eqs. (3.17). Thus we see that the approximation $w=-B/8$ may be applied without serious error. It is also noticed that, since B is no longer proportional to w, the limiting values of B and w do not take place simultaneously.

We shall next show that the stability limits are given by the limiting values of B or, in other words, that the parameters θ's in Hill's equation (3.5) take the characteristic values when $B=0$ or maximum.§§ For these limiting values of B, the values of y and w are calculated by Eqs. (3.17), and the parameters θ's are found from Eqs. (3.5). They are shown in Table I.

§§ In the case when the nondissipative system is considered, the characteristic exponent in the solution of Hill's equation becomes zero at the stability limits.

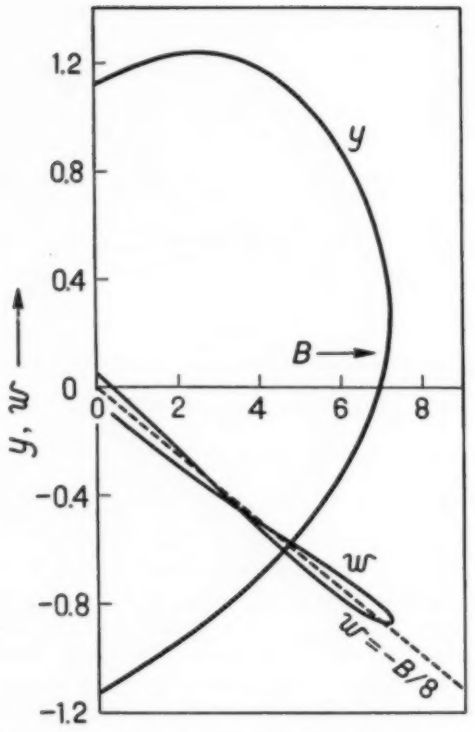


FIG. 5. Relationship between y, w, and B in Eqs. (3.17).

The boundaries of the first unstable region are given by Eqs. (3.6), namely,

$$\theta_0 = 1 \pm \theta_1. \tag{3.18}$$

Substituting the values of θ_1 of Table I, we obtain

$$\theta_0 = 2.0392 \quad \text{for} \quad B = 0,$$
$$\theta_0 = 1.3078 \quad \text{for} \quad B = \max.$$

These values differ from those given in Table I by 6.8 percent and 5.1 percent, respectively. This is due to the deficiency of approximation in Eq. (3.18). Hence, for the accurate values of y, w calculated by Eqs. (3.17), we have to apply the closer approximation given, for example, by the development[4]

$$\theta_0 = 1 \pm \theta_1 - \frac{1}{8}\theta_1^2 - \frac{1}{6}\theta_2^2 - \frac{1}{16}\theta_3^2 \pm \frac{1}{4}\theta_1\theta_2 \pm \frac{1}{12}\theta_2\theta_3$$
$$\mp \frac{1}{64}\theta_1^3 + \frac{1}{48}\theta_1^2\theta_2 \pm \frac{5}{192}\theta_1^2\theta_3 \pm \frac{1}{48}\theta_2^2\theta_3$$
$$\mp \frac{1}{144}\theta_1\theta_2^2 \mp \frac{1}{2304}\theta_1\theta_3^2 - \frac{13}{288}\theta_1\theta_2\theta_3 + \cdots. \tag{3.19}$$

Now, substituting the values of $\theta_1 \sim \theta_3$ of Table I into Eq. (3.19), we obtain

$$\theta_0 = 1.9038 \quad \text{for} \quad B = 0,$$
$$\theta_0 = 1.2352 \quad \text{for} \quad B = \max,$$

which differ from the values given in Table I by 0.28 percent and 0.73 percent respectively. Thus the discrepancies are considerably reduced, and we may conclude that the stability limits are given by the condition that the applied force B has its limiting values.‖‖

4. SUBHARMONIC OSCILLATION OF ORDER 1/3 WITH THE NONLINEAR CHARACTERISTIC $u = c_1 v + c_5 v^5$

For the brevity of calculation, we consider the non-dissipative case only. Putting

$$\left.\begin{aligned} &\nu = 3, \quad \delta = 0, \\ &c_2 = c_3 = c_4 = c_6 = \cdots = 0, \end{aligned}\right\}$$

in Eqs. (1.1) and (2.1), and remembering the condition (2.2), we have

$$\frac{d^2v}{d\tau^2} + v = c_5(v - v^5) + B\cos 3\tau. \tag{4.1}$$

Substituting the periodic solution

$$v = y\cos\tau + w\cos 3\tau, \quad w = -B/8,$$

into Eq. (4.1), and equating to zero the coefficient of $\cos\tau$, we obtain

$$y^4 + \frac{5}{2}y^2w + 6y^2w^2 + 3yw^3 + 3w^4 - \frac{8}{5} = 0. \tag{4.2}$$

The relationship between y and w is illustrated in Fig 6.

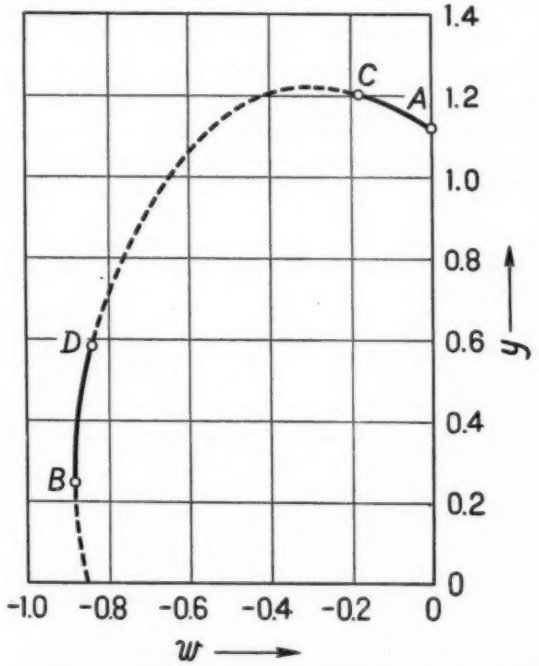


FIG. 6. Amplitude characteristic of the ⅓-harmonic oscillation (nonlinearity by quintic function).

[4] E. L. Ince, Monthly Notices Roy. Astron. Soc. **75**, 436 (1915).

‖‖ A more complete discussion on the stability limits will be reported later.

Negative part of y is omitted in the figure, because, as will be shown later, the equilibrium states in this part are unstable.

The stability problem may be treated in the same manner as in the preceding section, and the variational equation leads to

$$\left.\begin{aligned}&\frac{d^2\xi}{d\tau^2}+[\theta_0+2\sum_{s=1}^{6}\theta_s\cos 2s\tau]\xi=0,\\ &\text{with}\\ &\theta_0=c_1+\frac{5}{8}c_5(3y^4+4y^3w+12y^2w^2+3w^4),\\ &\theta_1=\frac{5}{4}c_5(y^4+3y^3w+3y^2w^2+3yw^3),\\ &\theta_2=\frac{5}{16}c_5(y^4+12y^3w+6y^2w^2+12yw^3),\\ &\theta_3=\frac{5}{4}c_5(y^3w+3y^2w^2+w^4),\\ &\theta_4=\frac{5}{8}c_5(3y^2w^2+2yw^3),\\ &\theta_5=\frac{5}{4}c_5yw^3,\\ &\theta_6=\frac{5}{16}c_5w^4.\end{aligned}\right\}\quad(4.3)$$

The stability condition for $n=1$ is obtained by substituting θ_0 and θ_1 into Eq. (3.6). Further by virtue of Eq. (4.2), we ultimately find

$$y^3+\frac{15}{8}y^2w+3yw^2+\frac{3}{4}w^3>0,\quad\text{for}\quad w<0.\qquad(4.4)$$

Referring to Fig. 6, the equilibrium states in the interval between A and B satisfy the condition (4.4). The stability limits A and B are given by the conditions $w=0$ and $w=\max$, respectively.

Figure 7 shows the loci of $\theta_n(n=1,2,3)$ which are drawn by varying the value of w (or B) for the limiting case of $c_1=0$ and $c_5=1$. As expected from the stability condition for $n=1$, θ_1 enters into the first unstable region in the dotted-line interval ab. We see moreover that θ_2 enters into the second unstable region in the dotted-line interval cd. Hence the stability condition

TABLE I. Values of y, w, and the θ's for $B=0$ and $B=\max$.

B	y	w	θ_0	θ_1	θ_2	θ_3
0	1.127	0.0510	1.9092	1.0392	0.0862	0.0019
Max = 7.205	0.284	−0.8655	1.2443	−0.3078	−0.3681	0.5619

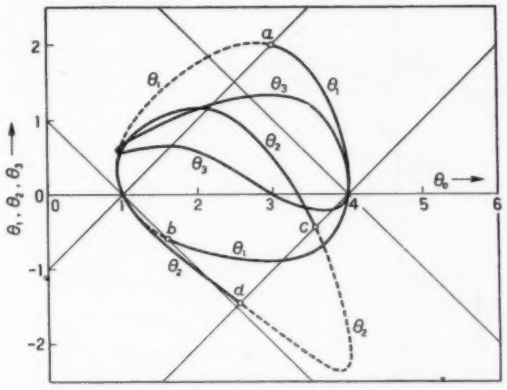


FIG. 7. Loci of θ's in Eqs. (4.3) with varying w.

for $n=2$ is no more satisfied in the interval cd, and the oscillation of order $\frac{2}{3}$ is excited, thus disturbing the continuation of the original subharmonic oscillation.

As mentioned above, the curves in Fig. 7 are drawn for the case of $c_1=0$ and $c_5=1$. With increasing c_1 (or decreasing c_5), however, these curves move towards the point (1,0), as we see from the expressions for the θ's in Eqs. (4.3). Hence, as the departure from linearity is reduced, the interval cd in the second unstable region contracts and finally disappears. We also see that in no case the parameters $\theta_3\sim\theta_6$ enter into their corresponding unstable regions. Hence, in the end, it may be concluded that the stability conditions for $n=2$ as well as for $n=1$ must be considered in the case when the nonlinearity is characterized by a quintic function. The points A, B, C, D in Fig. 6 correspond, respectively, to the critical points a, b, c, d in Fig. 7, so that the subharmonic oscillation of order $\frac{1}{3}$ is maintained in the stable state only in the intervals AC and BD.

We shall now discuss the approximation in the foregoing analysis just as we did at the end of the preceding section. Let the differential equation be given by

$$\frac{d^2v}{d\tau^2}+v^5=B\cos 3\tau.$$

Substituting the periodic solution

$$v=y\cos\tau+w\cos 3\tau,$$

and equating to zero the coefficients of $\cos\tau$ and $\cos 3\tau$ separately, we obtain

$$\left.\begin{aligned}&y^4+\frac{5}{2}y^2w+6y^2w^2+3yw^3+3w^4-\frac{8}{5}=0,\\ &\frac{5}{16}y^5+\frac{15}{8}y^4w+\frac{15}{8}y^3w^2+\frac{15}{4}y^2w^3+\frac{5}{8}w^5-9w=B.\end{aligned}\right\}\quad(4.5)$$

Figure 8 is obtained by plotting Eqs. (4.5). Thus we see that the approximation $w=-B/8$ may also be applied without serious error.

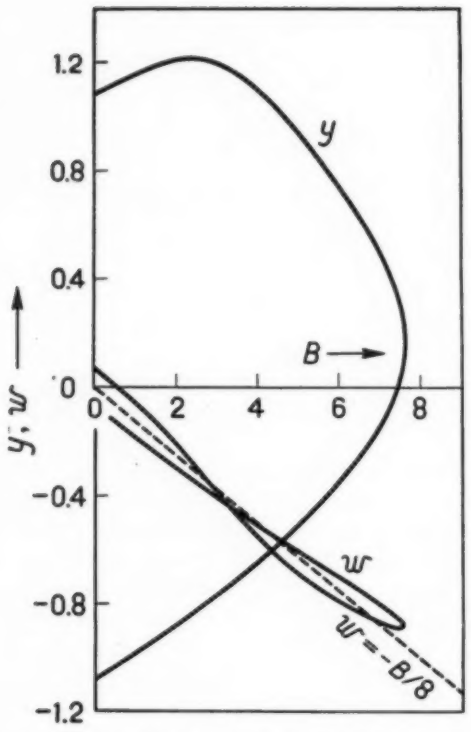


FIG. 8. Relationship between y, w, and B in Eqs. (4.5).

For $B=0$ and $B=\max$, the values of y, w and the parameters $\theta_0 \sim \theta_3$ in Eqs. (4.3) are calculated and shown in Table II.

Substituting the values of θ_1 into Eq. (3.18), we have

$$\theta_0=3.0327 \quad \text{for} \quad B=0,$$
$$\theta_0=1.3385 \quad \text{for} \quad B=\max.$$

These values differ from those given in Table II by 9.1 percent and 5.8 percent, respectively. Applying again the closer approximation (3.19) instead of Eq. (3.18), we obtain

$$\theta_0=2.7234 \quad \text{for} \quad B=0,$$
$$\theta_0=1.2415 \quad \text{for} \quad B=\max.$$

The discrepancies are reduced down to 1.9 percent and 2.0 percent, respectively, and we may conclude that the stability limits of the first unstable region are given by the condition that the applied force B has its limiting values.

We have so far discussed the stability limits of the first unstable region. A similar investigation may be carried out for the second unstable region. But in this case the stability limits have no particular relation to the applied force, and besides, as we shall see in a later experiment, a small parasitic oscillation of order $\frac{2}{3}$ can exist along with the original subharmonic oscillation in the neighborhood of the stability limits, so that the points c, d in Fig. 7, however accurately they might be determined, would not represent the exact critical points at which the original subharmonic oscillation ceases. Therefore, a further investigation into the stability limits of the second unstable region is not important and is omitted here.

TABLE II. Values of y, w, and the θ's for $B=0$ and $B=\max$.

B	y	w	θ_0	θ_1	θ_2	θ_3
0	1.0759	0.0716	2.7795	2.0327	0.7655	0.1337
Max=7.546	0.1542	−0.8818	1.2651	−0.3385	−0.3736	0.8210

5. EXPERIMENTAL CONSIDERATION

In the present section we shall compare the theoretical results (in the preceding sections) with experiments conducted for an electrical oscillatory circuit containing a saturable iron-core inductance and a capacitance. As shown previously by the author,[1] the circuit equation has the form of (1.1) when an alternating voltage (60 cps in our case) is applied to the circuit. When an appropriate initial condition is prescribed, the subharmonic oscillation of order $\frac{1}{3}$, i.e., of 20 cps, may easily be started in the circuit.

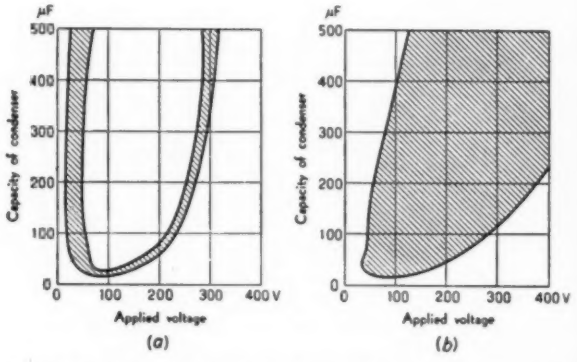


FIG. 9. Regions in which the $\frac{1}{3}$-harmonic oscillation is sustained. a. Magnetization curve by Eq. (5.1). b. Magnetization curve by Eq. (5.2).

Now, making use of a transformer-core inductance as the nonlinear element, we first determine the region in which the subharmonic oscillation of order $\frac{1}{3}$ is sustained. In Fig. 9a this region is depicted by hatched lines. The appearance of the blank part inside the sustaining region is an arresting feature and was previously reported by the present author[5] and others.[6] But no theoretical consideration was given at that time.

From the preceding analysis, however, it will be deduced that the blank part mentioned above corresponds to the unstable regions of order $n \geqq 2$, because the nonlinear characteristic, i.e., the saturation curve, of the ordinary transformer core is expressed by

$$f(v)=c_1v+c_3v^3+c_5v^5+c_7v^7+\cdots, \qquad (5.1)$$

where v represents the magnetic flux and $f(v)$ the magnetizing current, and the coefficients c_5, c_7, $\cdots$, usually

[5] C. Hayashi, Mitsubishi Denki 18, 128 (1942) [in Japanese].
[6] For instance, J. D. McCrumm, Trans. Am. Inst. Elec. Engrs. 60, 533 (1941).

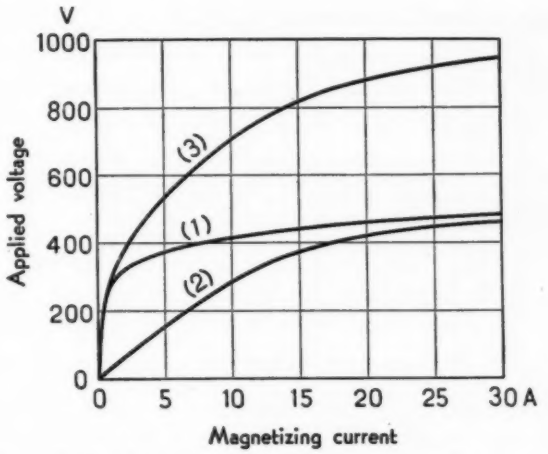


FIG. 10. Combined characteristic approximated to a cubic curve; (1) without air gap, (2) with air gap, (3) combined characteristic of (1) and (2).

predominate over c_1, c_3. If we use a core whose characteristic is expressed by

$$f(v)=c_1v+c_3v^3, \tag{5.2}$$

the blank part inside the sustaining region will be eliminated. Such a core is not available in practice, but we can obtain the characteristic of (5.2) by connecting a number of inductance coils in series and adjusting the length of the air gap which is interposed in each core. Figure 10 shows an example in which two cores are used, one with air gap and the other without. The resultant characteristic shows a fairly good approximation to Eq. (5.2). By making use of this composite inductance, the sustaining region of the subharmonic oscillation of order $\frac{1}{3}$ is determined and plotted in Fig. 9b. We see that the unstable oscillations corresponding to the unstable regions of order $n\geqq 2$ are completely excluded, and so the experimental verification is quite satisfactory.

We further measure the harmonic content in these oscillations with a heterodyne harmonic analyzer. Figure 11a, b shows the result for the cases in which the

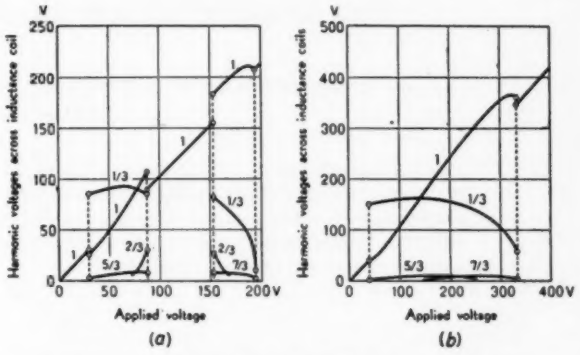


FIG. 11. Harmonic analysis of oscillations. a. Magnetization curve by Eq. (5.1). b. Magnetization curve by Eq. (5.2).

nonlinearities are given by Eqs. (5.1) and (5.2), respectively. In Fig. 11a, we observe the higher harmonics of orders 2/3, 5/3, 7/3, $\cdots$, among which the oscillation of order 2/3 is significant, because it is this oscillation (related to the second unstable region) that grows up rapidly and interrupts the original subharmonic oscillation [see Figs. 7 and 9a]. Whereas, in Fig. 11b, no such obstructive oscillation is observed, and the subharmonic oscillation of order $\frac{1}{3}$ is sustained in the whole region [see Figs. 2 and 9b].

Finally it is added that the subharmonic oscillation of order $\frac{1}{3}$ can occur when the nonlinearity is given by Eq. (5.1), but this oscillation is by no means observed when the nonlinearity is given by Eq. (5.2). These results also agree with the investigation in Sec. 2.

ACKNOWLEDGMENTS

The author wishes particularly to acknowledge his indebtedness to Dr. R. Torikai at whose instigation this study was started. The author's thanks are also due to Dr. T. Matsumoto and Dr. S. Tomotika for their valuable criticisms and suggestions, and to Mr. H. Takatsuki for his assistance in carrying out the calculations. In publishing this paper, the author was fortunate in having the help of Dr. Tomotika.

Null Characteristics of the Rotating Adcock Antenna System

JAMES G. HOLBROOK
Electrical Engineering Department, Northrop Aircraft, Inc., Hawthorne, California
(Received October 6, 1952)

It has been the usual practice when designing a rotating Adcock antenna system to allow a spacing of one-half wavelength between opposite dipoles at the highest intended frequency of operation. A mathematical analysis of this antenna system is presented, and it is shown that an optimum spacing for the most clearly defined nonmultiple nulls approaches a full wavelength. Useful design formulas are developed and typical response curves are presented graphically.

I. INTRODUCTION

MODERN radio direction finding stations use antenna systems which range from simple loops to the large fixed Adcock arrays with either mechanical[1] or electronic[2] goniometers. The small loop becomes useless for very long range bearings, since the incoming signal consists entirely of sky-wave components which have completely random polarizations; while on the other hand, the elaborate fixed Adcock systems, although responsive only to the vertically polarized components of the incoming signal, suffer from effective changes in antenna mast spacing when the incoming signal arrives at a large angle of vertical incidence.

For these reasons many users have retained the rotating Adcock[3] as the primary standard of direction finding. Its only serious operational disadvantage is the relatively long time that is necessary to properly evaluate the null point, and since the rotating array is the only type which is free from both polarization and vertical incidence errors, it seems destined to at least supplement the larger direction finding systems for years to come. The primary purpose of this paper is to discuss the null characteristics and optimum spacing of the dipoles.

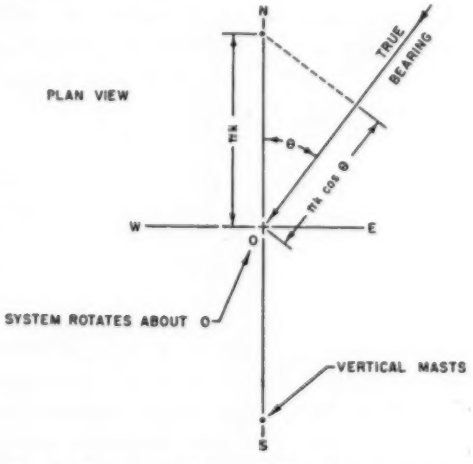


FIG. 1. Geometry of the rotating Adcock system.

[1] R. Keen, *Wireless Direction Finding* (Iliffe and Sons, Ltd., London, 1938).
[2] Paul G. Hansel, Electronics **21**, 86 (1948).
[3] Donald S. Bond, *Radio Direction Finders* (McGraw-Hill Book Company, Inc., New York, 1944).

II. SYMBOLS

E_O = Reference voltage measured at origin in Fig. 1.
E_N = Complex voltage measured at north antenna (Fig. 1).
E_S = Complex voltage measured at south antenna (Fig. 1).
E_{N-S} = Voltage fed from the Adcock system into receiver.
θ = Azimuth angle between $N-S$ antennas and coming signal (Fig. 1).
k = Antenna spacing to wavelength ratio, physical.
k' = Antenna spacing to wavelength ratio, effective.
ψ = Angle of vertical incidence of incoming wave, measured from horizontal.
α = Null width, defined in text.

III. THE GENERAL EQUATION

Consider a set of vertical dipoles located in the north-south plane as shown in the plan view of Fig. 1. The antennas are mounted normally upon a beam which is free to rotate about a center O. With respect to the origin O, the voltage at the north antenna may be written

$$E_N = E_O \epsilon^{j\pi k \cos\theta},$$

and at the south antenna

$$E_S = E_O \epsilon^{-j\pi k \cos\theta}.$$

The combined output voltage is the difference, rather than the sum, because of the conventional Adcock transposition in the connecting transmission line, and the actual output from antenna to receiver is

$$E_{N-S} = E_O(\epsilon^{j\pi k \cos\theta} - \epsilon^{-j\pi k \cos\theta})$$

or, factoring out a $2j$,

$$E_{N-S} = 2jE_O \sin(\pi k \cos\theta). \quad (1)$$

This may be normalized by letting $2jE_O = 1$, whereupon the general equation becomes

$$E_{N-S} = \sin(\pi k \cos\theta). \quad (1a)$$

A plot showing the general trend of this equation for two values of k is shown in Fig. 2.

It will be noted immediately that for small values of k, the maximum output will occur when $\theta = 0$. At this

point

$$E_{N-S}=\sin(\pi k). \tag{2}$$

For large values of k, the maximum output will no longer occur at $\theta=0$, but at

$$\theta=\cos^{-1}\left[\frac{N}{2k}\right], \quad \text{where} \quad N=1, 3, 5, 7\cdots\leqslant 2k. \tag{3}$$

For purposes of comparison all maximum values may be plotted as relative values of 1.0, as in Fig. 2. From Eq. (1a) nulls will occur when

$$\pi k\cos\theta=M\pi$$

or when

$$\theta=\cos^{-1}\left[\frac{M}{k}\right], \quad \text{where} \quad M=0, 1, 2, 3\cdots. \tag{4}$$

IV. MULTIPLE NULLS

Multiple nulls are defined here as more than two nulls occurring in the 360 degrees of azimuth. Before making numerical comparisons we shall proceed as follows: Taking the derivative of Eq. (1a),

$$\frac{dE_{N-S}}{d\theta}=-\pi k\cos(\pi k\cos\theta)\sin\theta, \tag{5}$$

and setting equal to zero,

$$\pi k\cos(\pi k\cos\theta)\sin\theta=0. \tag{6}$$

It is seen that the slope is always O for $\theta=0$ and, in addition, will also be O when

$$\theta=\cos^{-1}[N/2k] \quad N=1, 3, 5, 7\cdots\leqslant 2k. \tag{3a}$$

Equation (3a) will have real solutions only when $N/2k$ lies between 0 and 1, for $N=1$ (the normal mode of operation), k must be greater than $\frac{1}{2}$ for Eq. (3a) to have real solutions. It is to be noted, however, that real solutions for Eq. (3a) does not necessarily mean that multiple nulls will occur. These will not occur unless there are also solutions to Eq. (4) when $M=1$, or, multiple nulls cannot occur unless $k\geqslant 1$.

V. WIDTH OF NULLS

Any definition of null width is necessarily arbitrary. In this paper we shall define the width of nulls (α) as the angle between two points which are equally spaced about the null point, and where the amplitude is down to 50 percent of the peak value. It is seen from Eq. (1a) that for very small values of k the sine of the angle becomes equal to the angle, and the response curve is then merely a cosine wave. By the above definition then, the null width α becomes 60 degrees. This represents the poorest possible null. Larger values of k may be evaluated from Eq. (1a) by letting

$$\sin(\pi k\cos\theta)=\sin(\pi/6)$$

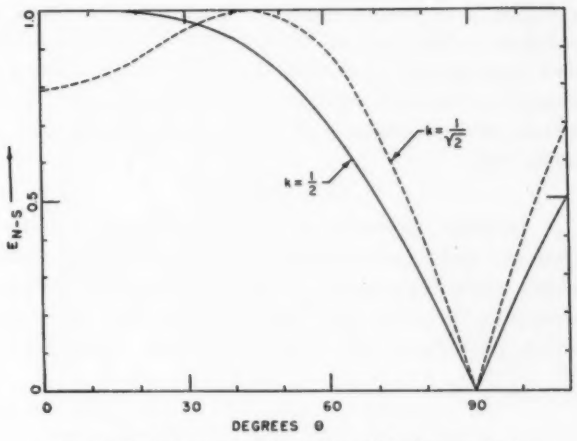


FIG. 2. Variations of null quality with changes in k.

or

$$\theta=\cos^{-1}(1/6k) \tag{7}$$

so that

$$\alpha=\pi-2\cos^{-1}(1/6k), \tag{8}$$

or the null width is given as a function only of k. It is seen from Eqs. (2) and (4) that k must be less than one in order to avoid multiple nulls; however, the value of $k=0.9$ is entirely satisfactory, and by using Eq. (8) it is found that $\alpha=21$ degrees. This null is therefore about three times as sharp or well defined, as the null obtained from Eq. (1a) when k is very small. In addition, the larger value of k provides increased sensitivity at other than null points.

At the actual null point (90° in Fig. 2) it is found that Eq. (5) reduces to

$$\left.\frac{dE_{N-S}}{d\theta}\right|_{\theta=\pi/2}=-\pi k. \tag{9}$$

By means of this expression the nulls may be compared for quality on the basis of steepness of slope rather than null width, it being understood that nulls having the steepest slopes are the most clearly defined. Taking as a limit $k=1$ gives a slope of 3.14+, which is the maximum slope without encountering multiple nulls. If there is sufficient time to evaluate and resolve multiple nulls, the true nulls may be made extremely sharp by choosing

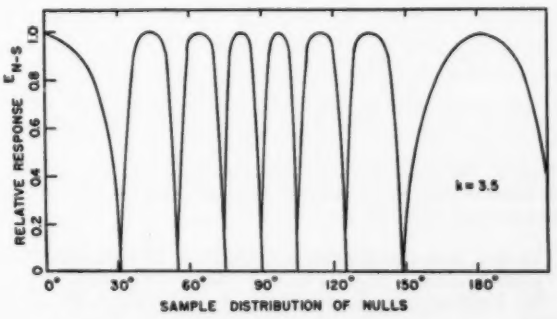


FIG. 3. Nonuniform spacing of multiple nulls.

a value of k greater than 1. That such a process is possible is clear from an inspection of Fig. 3, where it is seen that the true null at 90 degrees is easily isolated as a result of the spacing symmetry. This figure is shown merely as an extreme case and is not recommended for actual use.

VI. CONCLUSIONS

A possible objection to the validity of letting k approach unity may arise because of the secondary nulls which would then arise at 0 and 180 degrees. In actual operation, however, this would require that the signal arrive parallel to the earth's surface. Most signals arrive from some vertical angle and usually the secondary nulls will be automatically destroyed by the reduction in effective k, k' given as

$$k' = k \cos\psi, \tag{10}$$

where ψ is the vertical angle of arrival of the wave as measured from horizontal.

From the foregoing discussion it is shown that a manually rotated Adcock system will produce most clearly defined nulls when k is large for the highest frequency of operation. The optimum value appears to be only slightly less than one.

JOURNAL OF APPLIED PHYSICS VOLUME 24, NUMBER 5 MAY, 1953

A Method for the Quantitative Determination of Preferred Orientation

L. K. JETTER AND B. S. BORIE, JR.
Metallurgy Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee
(Received October 2, 1952)

A new method for the study of preferred orientation in polycrystalline materials with a Geiger counter x-ray spectrometer is described. The unique feature of the method is that a spherical diffraction specimen is employed. Sufficient data for the construction of a complete pole figure may be obtained, and the observed intensities may be plotted directly on a stereographic net without geometry or absorption corrections. The method may be used satisfactorily with the heavy metals since x-rays are not transmitted through the specimen. A technique for accurately machining and electroetching spherical specimens is given, and the special goniometer used with the spectrometer is described. Sample data are given and their interpretation discussed. The method is compared with other recently developed techniques for the study of preferred orientation.

INTRODUCTION

THE quantitative study of preferred orientation has been made practical by the recent development of several Geiger counter x-ray spectrometer techniques. Among the new methods are those of Decker, Asp, and Harker;[1] Norton;[2] Field and Merchant;[3] and Schulz.[4] Each employs a special specimen holder mounted on a conventional focusing spectrometer. The specimen orientation is varied while the Geiger counter, held stationary at the appropriate Bragg angle, measures an intensity which is related to the fraction of the irradiated volume of the specimen properly oriented to diffract.

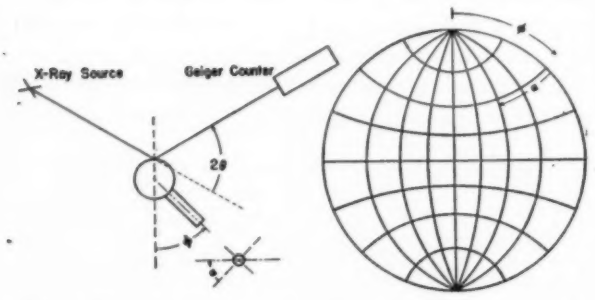


FIG. 1. Diagram showing diffraction geometry for this method and relationship of specimen orientation to position of diffracting plane normal on Wulff net.

With the exception of the technique described by Norton, flat specimens are employed. This requires either that the primary beam intensity be severely limited[4] or that the diffracted beam intensity be corrected for variation in diffraction geometry with specimen orientation.[1,3] Norton's method is characterized by the use of a cylindrical specimen. For orientation changes by rotation about the cylinder axis, no intensity correction is necessary.

With any of these methods, construction of complete pole figures requires the use of a number of identical specimens cut at various angles to some principal direction of the sample.

The unique feature of the method to be described here is that a spherical diffraction specimen is employed. For orientation changes by rotation about any axis through the center of the sphere, the diffraction geometry is not altered; hence, no intensity correction is required. Complete pole figure data may be obtained with only one specimen; and, since the x-ray beam is re-

[1] Decker, Asp, and Harker, J. Appl. Phys. **19**, 388–392 (1948).
[2] J. T. Norton, J. Appl. Phys. **19**, 1176–1178 (1948).
[3] M. Field and M. E. Merchant, J. Appl. Phys. **20**, 741–745 (1949).
[4] L. G. Schulz, J. Appl. Phys. **20**, 1030–1033 (1949).

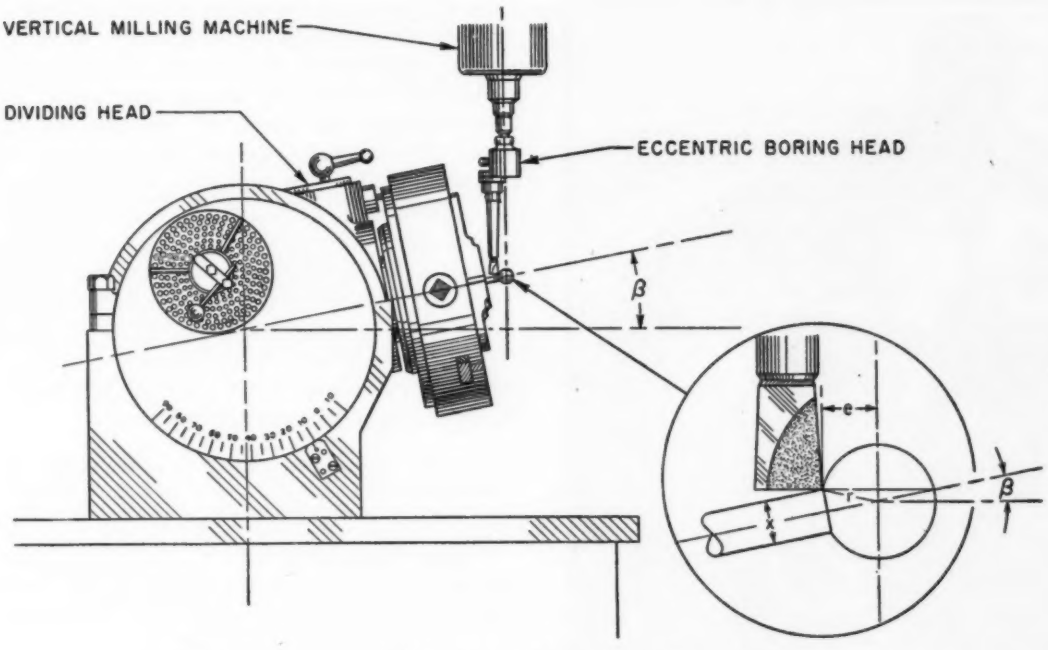


FIG. 2. Arrangement for machining spherical x-ray diffraction specimen.

flected from rather than transmitted through the specimen, the method may be used satisfactorily regardless of the absorption coefficient of the material to be examined.

PRINCIPLE OF THE METHOD

Appended to the spherical diffraction specimen is a cylindrical stem by which the specimen is mounted on the spectrometer. The stem axis is used to define some principal direction of the sample from which the specimen is taken; reference marks on the stem surface may be used to identify other directions of importance. By means of a special goniometer the specimen is mounted on the spectrometer so that the stem axis is in the plane of the incident and the diffracted beams—the plane of the spectrometer. The goniometer provides for rotation of the sphere about the stem axis through an angle α and also for rotation of the sphere about an axis normal to the spectrometer plane through an angle ϕ. Both axes pass through the center of the sphere. These motions of the specimen relative to the spectrometer plane are illustrated in Fig. 1. When the angle ϕ is zero, the stem axis bisects the angle between the incident and diffracted beams, and only that part of the irradiated volume oriented so that the plane normal is parallel to the stem axis will diffract. If the position of the diffracting plane normal for this specimen orientation is represented by the pole of a Wulff net, then the position of the diffracting plane normal for any other specimen orientation is represented by the coordinates ϕ and α, as shown in Fig. 1.

It is clear that after any rotation through ϕ and α the measured diffraction intensity is directly comparable to that measured at $\phi=0$ since the diffraction geometry is not altered. It is also evident that the only limitation on the value of ϕ is that it not be so large that the specimen stem intercepts the diffracted beam. Since this maximum ϕ is always greater than 90° and since the rotation through α is unlimited, complete data for construction of the pole figure may be obtained directly.

PREPARATION OF THE SPECIMEN

An accurately spherical diffraction specimen may be machined on a vertical milling machine equipped with an eccentric boring tool and a dividing head, as shown in Fig. 2. The dividing head must be positioned so that its rotation axis intersects that of the boring head. The angle of inclination β of the dividing head and the eccentricity e of the boring head are related to the specimen radius r and the stem diameter x as follows:

$$\beta=\tfrac{1}{2}\arcsin(x/2r),$$
$$e=r\cos\beta.$$

The stem of desired diameter is machined on a lathe. By means of this stem the specimen is mounted in the dividing head of the milling machine. The sphere is cut by rapidly rotating the eccentric cutter while slowly rotating the dividing head and slowly elevating the milling machine table. When the cutter just contacts the stem, the specimen is fully formed. The machined sphere is lapped, using a geometrically similar arrangement in which a tube of appropriate diameter replaces the eccentric cutter, and is then etched to remove the surface layer deformed by machining and lapping.

A metal specimen may be electroetched with the aid of a cathode in the form of a hollow sphere. To maintain

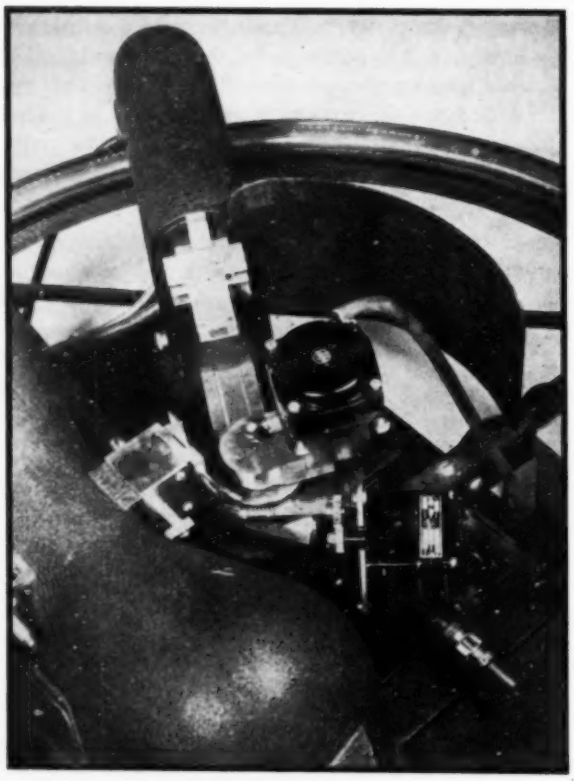


FIG. 3. Photograph of special goniometer with spherical diffraction specimen mounted on horizontal Geiger counter x-ray spectrometer.

the sphericity of the specimen, conditions are adjusted so that etching occurs slowly until nearly the desired amount of metal is removed; then etching is speeded up to remove pitting and polish the sphere. With this technique, specimens as small as 0.2 inch in diameter and spherical to within 0.001 inch are consistently obtained.

Even for extremely thin sections, this method may still be employed. For instance, a hemispherical end may be lapped on the end of fine wire. For thin sheets, a composite sample of many sections identically oriented may be made up and then machined as described above.

DESCRIPTION OF THE EQUIPMENT

The special goniometer together with the diffraction specimen is shown mounted on a horizontal Geiger counter x-ray spectrometer in Fig. 3. The goniometer consists of a duplex base, the upper section of which is dovetailed into the lower and may be moved linearly with respect to the lower section by means of a micrometer screw. The lower section is fastened to the goniometer bushing of the spectrometer. The upper section supports a carriage which may be rotated about a vertical axis; this carriage in turn supports a horizontal spindle to which the specimen is attached by means of a chuck. The two axes intersect. The specimen is positioned so that the center of the sphere lies on the vertical rotation axis of the carriage. A horizontally mounted Bodine synchronous motor with accessory gearing provides for automatic rotation of the carriage about its vertical axis through the angle ϕ. Another Bodine motor, mounted vertically, provides for automatic rotation of the spindle about its horizontal axis through the angle α. Appropriate dials permit measurements of these angles. By means of an eccentric, the latter motor also provides for oscillatory motion superimposed on the rotation through ϕ. This motion affords better statistics when coarse-grained specimens are being examined.

Thus, preliminary to measurement of the preferred orientation, the specimen may be subjected to three independent motions for its alignment:

1. It rotates through an angle θ about the vertical spectrometer axis as the Geiger counter arm is rotated through 2θ, the scattering angle associated with the family of crystallographic planes to be studied.
2. Sliding the specimen in the chuck moves it linearly along the axis of the specimen stem. This motion permits the sphere to be positioned on the vertical rotation axis of the carriage.

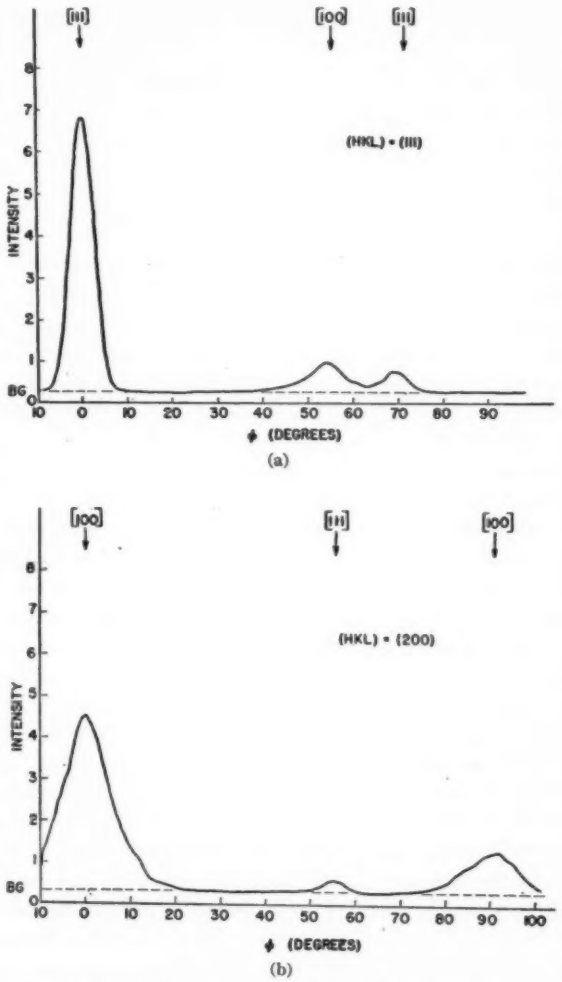


FIG. 4. Preferred orientation plots for the (111) and (200) planes of extruded aluminum.

3. The specimen is translated in the direction of the diffracting plane normal by rotation of the micrometer screw which slides the upper section of the base over the lower. This permits the sphere to be moved into the primary beam, aligning it for optimum diffraction geometry.

EXPERIMENTAL PROCEDURE

With the sphere aligned in the primary beam and oriented to give maximum pole density, the diffracted beam is scanned in the usual manner to establish the background intensity and the angle of maximum intensity for the crystallographic plane under study. The counter then is held stationary at this angle and the intensity changes measured on a potentiometer recorder as the sphere is rotated.

Several procedures may be employed for scanning the sphere through the required range of ϕ and α. With the sphere held constant at some α, it may be automatically rotated through ϕ, and then the procedure repeated for other values of α. Alternately, the sphere may be automatically and simultaneously rotated through both ϕ and α, the rates of rotation being so adjusted as to scan the sphere adequately, and all the data recorded on one chart. For specimens having a fiber texture, the fiber axis of which is parallel to the stem axis, the sphere may be rotated rapidly through α while it is rotated slowly through ϕ.

It has been found that the principal effect of small specimen misalignments is to cause small changes in the observed value of 2θ for different ϕ or α. Misalignment, if it is not large, seems to have a negligible effect on the observed intensity. Hence, whenever possible, wide slits are used at the Geiger counter, giving a broad, relatively slowly changing diffraction maximum. The observed count rate then is essentially independent of small shifts in 2θ caused by the misalignment. When such a wide counter slit is used, it has been found convenient to use a set of Soller slits at the counter to keep the background to a reasonably low level (see Fig. 3). Although the diffraction lines of cubic metals are usually spaced so that wide slits may be used, for metals of lower symmetry it is often necessary to use narrow slits to resolve a set of closely spaced lines. In the latter case, very careful specimen preparation and alignment are required.

TYPICAL RESULTS

Shown in Fig. 4 are plots of diffraction intensity *vs* ϕ obtained as described above on a 0.50-inch diameter spherical specimen machined from a 1.5-inch diameter extruded aluminum (99.9 percent) rod. The billet temperature was 480°C and the extrusion ratio 7.6. The specimen was machined so that the stem axis was parallel to the extrusion direction. Since no significant variation in texture with α is to be expected for extruded material, the fiber texture was determined by rotating slowly through ϕ while rotating rapidly through α for better statistics.

The characteristic [111]–[100] texture of deformed face-centered cubic metals is shown clearly by the two plots. For the (111) plot, the peaks observed at $\phi=0°$ and 70.5° are produced by the [111] component. The [100] component is evident from the peak at 54.7°. For the (200) plot, the peak at $\phi=54.7°$ is associated with the [111] component and those at $\phi=0°$ and 90° with the [100] component.

The two plots in Fig. 4 are tracings from the recorder chart. The specimen was aligned in the goniometer and the two records obtained in about one hour. For fiber textures of this type the significant data are clearly shown by these plots and actual pole figures are not needed.

ACKNOWLEDGMENT

The authors wish to thank Mr. R. O. Williams and Mr. G. T. Murray for their assistance in the development of the method and the Research Shops of the Oak Ridge National Laboratory for the development of the special machining technique.

Free Space Radiation Impedance of Rhombic Antenna*

JESSE GERALD CHANEY†
United States Naval Postgraduate School, Monterey, California
(Received October 13, 1952)

The expression for the driving point impedance of a generalized electric circuit as vizualized by the author is partially integrated, and the physical significance of certain terms is discussed in connection with their application to antenna problems. Upon postulating an unattenuated traveling wave as a first approximation to the current along either a terminated rhombic or Vee antenna, a formula is derived for the free space radiation impedance of each. The expression *radiation impedance* is used for that portion of the driving point impedance which may be determined from the complex power into the external fields. The resistive component of the impedance for the rhombic antenna checks with the radiation resistance as listed by Leonard Lewin in a discussion of a paper by Donald Foster, with Lewin apparently obtaining his formula by the solid angle Poynting vector method.

THE GENERALIZED CIRCUIT

IN the generalized circuit (Fig. 1), a slice generator is assumed to exist between terminals b and c, and the current is assumed continuous through the generator. Terminals d and e may or may not be closed.

The current along the circuit is assumed to be given by $I_0 f(P)$, with P being any position on the circuit, and with I_0 being the current through the generator. If the length of the circuit is l, the driving point impedance is given by[1]

$$Z_{in}=lZ_i f_m^2+j\frac{30}{k}\left(\oint_b^d+\oint_e^c\right)\left(\oint_b^d+\oint_e^c\right)\times Re[f(P_1)^* f(P_2)]\diamondsuit_1[e(r_{21})d\bar{r}_2]\cdot d\bar{r}_1, \quad (1)$$

in which

P_1 = any point along the axis of the wire,
P_2 = any point along the inner periphery of the wire,
r_{21} = distance from P_1 to P_2,
$e(r_{21})=r_{21}^{-1}\exp(-jkr_{21})$,
$k=\omega(\mu_0\epsilon_0)^{\frac{1}{2}}=2\pi/\lambda$,
$\mu_0=4\pi(10^{-7})$ henries per meter,
$\epsilon_0=(36\pi\cdot 10^9)^{-1}$ farads per meter,

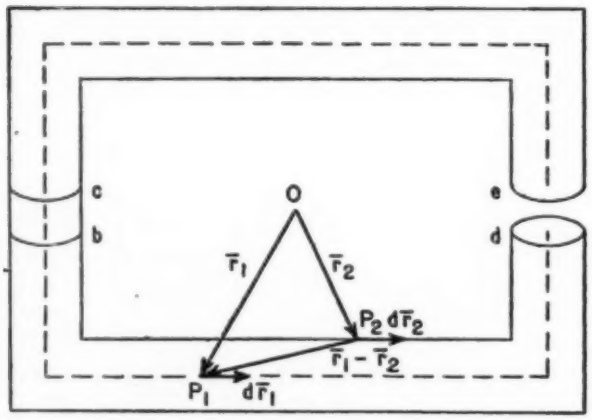


FIG. 1. The generalized circuit.

Z_i = internal impedance of the circuit per meter,
f_m^2 = mean square magnitude of the current distribution,
$\diamondsuit_1=\nabla_1\nabla_1\cdot+k^2$ = operator *deltil* or *delcap* with the subscript indicating the position at which the differentiations are to be performed,
* = the complex conjugate to be taken,
Re = the real part to be taken.

Now let A indicate the integrated value of the first term of the integrand, that is,

$$A=j\frac{30}{k}\left(\oint_b^d+\oint_e^c\right)\left(\oint_b^d+\oint_e^c\right)\times Re[f(P_1)^* f(P_2)]\nabla_1[\nabla_1\cdot e(r_{21})d\bar{r}_2]\cdot d\bar{r}_1. \quad (2)$$

Upon integrating by parts along the axis, let

$$U=Re[f(P_1)^* f(P_2)],$$
$$dU=\nabla_1 Re[f(P_1)^* f(P_2)]\cdot d\bar{r}_1,$$
$$dV=\nabla_1[\nabla_1\cdot e(r_{21})d\bar{r}_2]\cdot d\bar{r}_1,$$
$$V=\nabla_1\cdot e(r_{21})d\bar{r}_2=-\nabla_2 e(r_{21})\cdot d\bar{r}_2,$$

and get

$$A=j\frac{30}{k}\left(\oint_b^d+\oint_e^c\right)\{Re[f(P_b)^* f(P_2)]\nabla_2 e(r_{2b})-Re[f(P_c)^* f(P_2)]\nabla_2 e(r_{2c})+Re[f(P_e)^* f(P_2)]\nabla_2 e(r_{2e})-Re[f(P_d)^* f(P_2)]\nabla_2 e(r_{2d})\}\cdot d\bar{r}_2+j\frac{30}{k}\left(\oint_b^d+\oint_e^c\right)\left(\oint_b^d+\oint_e^c\right)\times\{\nabla_1 Re[f(P_1)^* f(P_2)]\cdot d\bar{r}_1\}\{\nabla_2 e(r_{21})\cdot d\bar{r}_2\}. \quad (3)$$

Let B represent the last term of Eq. (3), and integrate along the perimeter,

$$B=j\frac{30}{k}\left(\oint_b^d+\oint_e^c\right)\left(\oint_b^d+\oint_e^c\right)\times\{\nabla_1 Re[f(P_1)^* f(P_2)]\cdot d\bar{r}_1\}\nabla_2 e(r_{21})\cdot d\bar{r}_2. \quad (4)$$

* U. S. Naval Postgraduate School Technical Report No. 4.
† Professor of Electronics.
[1] J. G. Chaney, J. Appl. Phys. 22, 12, 1429 (1951).

Now let

$$U=\nabla_1 Re[f(P_1)^* f(P_2)]\cdot d\bar{r}_1,$$
$$dU=\nabla_2\{\nabla_1 Re[f(P_1)^* f(P_2)]\cdot d\bar{r}_1\}\cdot d\bar{r}_2,$$
$$dV=\nabla_2 e(r_{21})\cdot d\bar{r}_2,$$
$$V=e(r_{21}),$$

and get

$$B=j\frac{30}{k}\left(\oint_b^d+\oint_e^c\right)\{e(r_{d1})\nabla_1 Re[f(P_1)^* f(P_d)]$$
$$-e(r_{e1})\nabla_1 Re[f(P_1)^* f(P_e)]$$
$$+e(r_{c1})\nabla_1 Re[f(P_1)^* f(P_c)]$$
$$-e(r_{b1})\nabla_1 Re[f(P_1)^* f(P_b)]\}\cdot d\bar{r}_1$$
$$-j\frac{30}{k}\left(\oint_b^d+\oint_e^c\right)\left(\oint_b^d+\oint_e^c\right)$$
$$\times e(r_{21})\nabla_2\{\nabla_1 Re[f(P_1)^* f(P_2)]\cdot d\bar{r}_1\}\cdot d\bar{r}_2. \quad (5)$$

Recall that the current distribution is postulated to be the same along both paths of integration, and that the radius of the wire is postulated to be sufficiently small for the two paths to be interchangeable. Hence upon interchanging the paths in the single integral terms of Eq. (5), and upon substituting from Eq. (5) into Eq. (3), and subsequently into Eq. (1), the driving point impedance becomes

$$Z_{in}=lZ_i f_m^2+j\frac{30}{k}\left(\oint_b^d+\oint_e^c\right)\left(\oint_b^d+\oint_e^c\right)$$
$$\times e(r_{21})\{k^2 Re[f(P_1)^* f(P_2)]d\bar{r}_1$$
$$-\nabla_2(\nabla_1 Re[f(P_1)^* f(P_2)]\cdot d\bar{r}_1)\}\cdot d\bar{r}_2$$
$$+j\frac{30}{k}\left(\oint_b^d+\oint_e^c\right)\{e(r_{2d})\nabla_2 Re[f(P_d)^* f(P_2)]$$
$$-Re[f(P_d)^* f(P_2)]\nabla_2 e(r_{2d})$$
$$-e(r_{2e})\nabla_2 Re[f(P_e)^* f(P_2)]$$
$$+Re[f(P_e)^* f(P_2)]\nabla_2 e(r_{2e})$$
$$+e(r_{2c})\nabla_2 Re[f(P_c)^* f(P_2)]$$
$$-Re[f(P_c)^* f(P_2)]\nabla_2 e(r_{2c})$$
$$-e(r_{2b})\nabla_2 Re[f(P_b)^* f(P_2)]$$
$$+Re[f(P_b)^* f(P_2)]\nabla_2 e(r_{2b})\}\cdot d\bar{r}_2. \quad (6)$$

Integrating the terms in the last integral containing the gradient of the exponential terms, and assuming the radius of the wire is a with

$$ka\ll 1, \quad a\ll l,$$

one obtains

$$Z_{in}=lZ_i f_m^2+\frac{1}{j\omega 2\pi\epsilon_0}\{\tfrac{1}{2}e(a)[|f(P_d)|^2+|f(P_e)|^2+2]$$
$$-e(r_{bc})Re[f(P_b)^* f(P_c)]$$
$$-e(r_{de})Re[f(P_d)^* f(P_e)]$$
$$-e(r_{bd})Re[f(P_b)^* f(P_d)]$$
$$+e(r_{cd})Re[f(P_c)^* f(P_d)]$$
$$-e(r_{ce})Re[f(P_c)^* f(P_e)]$$
$$+e(r_{be})Re[f(P_b)^* f(P_e)]\}$$
$$+j\frac{60}{k}\left(\oint_b^d+\oint_e^c\right)\{e(r_{2d})\nabla_2 Re[f(P_d)^* f(P_2)]$$
$$-e(r_{2b})\nabla_2 Re[f(P_b)^* f(P_2)]$$
$$+e(r_{2c})\nabla_2 Re[f(P_c)^* f(P_2)]$$
$$-e(r_{2e})\nabla_2 Re[f(P_e)^* f(P_2)]\}\cdot d\bar{r}_2$$
$$+j\frac{30}{k}\left(\oint_b^d+\oint_e^c\right)\left(\oint_b^d+\oint_e^c\right)e(r_{21})$$
$$\times\{k^2 Re[f(P_1)^* f(P_2)]d\bar{r}_1$$
$$-\nabla_2(\nabla_1 Re[f(P_1)^* f(P_2)]\cdot d\bar{r}_1)\}\cdot d\bar{r}_2. \quad (7)$$

If the circuit is now thought of as an antenna that is symmetrically fed, that is, if the antenna is physically symmetrical with respect to the generator, the sense of the path from e to c may be reversed in the single integral terms of Eq. (7), with the interchanging of c and e and b and d, respectively, within the interval (ec). Also, after replacing $-d\bar{r}_2$ with $+d\bar{r}_2$ along with the corresponding interchange of limits for the interval (ec), the operator $+\nabla_2$ is replaced with the operator $-\nabla_2$. Hence, for the symmetrically fed antenna, the single integral terms may be doubled and integrated only from b to d. However, if the terminal spacings bc and de are each negligibly small with respect to a wavelength, the current may be assumed continuous through the terminals and hence the single integral terms of Eq. (7) then vanish.

The real part of the terms within the brace of Eq. (7) represent the reactive terms due to the end capacitances, and the imaginary part represents a correction to the component of the radiation resistance due to the distributed charges. The latter correction is necessary because the path is not closed.

In practice, either the current is supposed to vanish at the terminals de or a load impedance Z_0 is assumed to be inserted at the terminals to cause the current to take on the postulated distribution. Also, the input terminals are shunted at the generator. Thus, the real part of the terms within the brace either vanish or are usually

discarded. The imaginary terms also either vanish or may be disregarded provided the terminal spacings are quite small in comparison with a wavelength. Otherwise, if f_0 is the current magnitude at d and at e, it may be necessary to retain the following correction terms from Eq. (7),

$$60\left(\frac{\sin kr_{bc}}{kr_{bc}}-1\right)+60\left(\frac{\sin kr_{de}}{kr_{de}}-1\right)f_0^2,$$

in which it is assumed that the distances r_{bd} and r_{be} do not appreciably differ from the distances r_{cd} and r_{ce}, respectively.

Thus, assuming the terminal spacings permit the dropping of the correction terms and that the current is continuous through the terminals so that the single integral terms in Eq. (7) vanish, the driving point impedance becomes,[1]

$$Z_{in}=lZ_i f_m^2+Z_0 f_0^2+j\frac{30}{k}\oint_1\oint_2 e(r_{21})\times\{k^2 Re[f(P_1)^* f(P_2)]d\bar{r}_1-\nabla_2(\nabla_1 Re[f(P_1)^* f(P_2)]\cdot d\bar{r}_1)\}\cdot d\bar{r}_2. \quad (8)$$

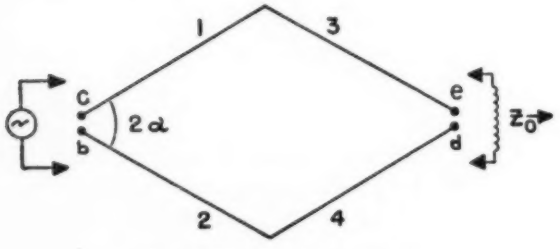


FIG. 2. The rhombic antenna.

In many cases of antenna applications, the current may be approximated sufficiently well by a distribution function which either satisfies or may be broken into the sum of functions which satisfy one of the following equations of constraint,

$$\nabla_2(\nabla_1 Re[f(P_1)^* f(P_2)]\cdot d\bar{r}_1)\cdot d\bar{r}_2 \pm k^2 Re[f(P_1)^* f(P_2)]dr_1 dr_2=0. \quad (9)$$

Substituting from Eq. (9) into Eq. (8),

$$Z_{in}=lZ_i f_m^2+Z_0 f_0^2+j\frac{\omega\mu_0}{4\pi}\oint_1\oint_2 e(r_{21})Re[f(P_1)^* f(P_2)]d\bar{r}_1\cdot d\bar{r}_2 \pm j\frac{\omega\mu_0}{4\pi}\oint_1\oint_2 e(r_{21})Re[f(P_1)^* f(P_2)]dr_1 dr_2. \quad (10)$$

The first integral is the generalized Neumann's formula, and is sometimes used alone for estimating the radiation resistance of a circuit. It appears that the second double integral term should not be neglected. However, it should be remembered that a constant current distribution does not satisfy Eq. (9), and that Eq. (10) does not contradict Neumann's formula *per se*.

If the terminals of Z_0 are sufficiently near each other for the closing of the integrals, Eq. (10) may be written

$$Z_{in}=lZ_i f_m^2+f_0^2 Z_0+j30k\oint_1\oint_2 e(r_{21})\times Re[f(P_1)^* f(P_2)](d\bar{r}_1\cdot d\bar{r}_2\pm dr_1 dr_2). \quad (11)$$

In terms of arc lengths, let

$$g(ks_1, ks_2)=Re[f(P_1)^* f(P_2)],$$
$$h(ks_1, ks_2)=e(r_{21}),$$
$$\cos\theta(s_1, s_2)ds_1 ds_2=d\bar{r}_1\cdot d\bar{r}_2,$$

and Eqs. (9) and (11) become, respectively,

$$\left(\frac{\partial^2}{\partial s_1\partial s_2}\pm k^2\right)g(ks_1, ks_2)=0; \quad (12)$$

$$Z_{in}=lZ_i f_m^2+Z_0 f_0^2+j30k\oint_1\oint_2 g(ks_1, ks_2)\times h(ks_1, ks_2)[\cos\theta(s_1, s_2)\pm 1]ds_1 ds_2. \quad (13)$$

If the algebraic sign within the parenthesis of Eq. (12) is positive,

$$Z_{in}=lZ_i f_m^2+Z_0 f_0^2+j60k\oint_1\oint_2 g(ks_1, ks_2)\times h(ks_1, ks_2)\cos^2\tfrac{1}{2}\theta(s_1, s_2)ds_1 ds_2, \quad (14)$$

and if it is negative,

$$Z_{in}=lZ_i f_m^2+Z_0 f_0^2-j60k\oint_1\oint_2 g(ks_1, ks_2)\times h(ks_1, ks_2)\sin^2\tfrac{1}{2}\theta(s_1, s_2)ds_1 ds_2. \quad (15)$$

THE RHOMBIC ANTENNA

For the purpose of obtaining a formula for the radiation impedance of a rhombic antenna in free space (Fig. 2), the generator is assumed to short-circuit terminals b and c, and a terminal impedance Z_0 is postulated at terminals d and e which, to a first approximation, causes the current to be an unattenuated travelling wave with no reflection. Each leg of the rhombic is assumed to be l meters in length, and the vertex angle at the generator is assumed to be 2α.

Upon selecting the origin at the generator, the variable s ranges from $-2l$ at e to $+2l$ at d. For s_1 and s_2 along different but parallel wires, $\theta=\pi$. For s_1 and s_2 along different wires *one* and *two* or *three* and *four*, $\theta=\pi-2\alpha$. For s_1 and s_2 along different wires *two* and *three* or *four* and *one*, $\theta=2\alpha$.

For negative s, $I=I_0e^{jks}$, and for positive s, $I=I_0e^{-jks}$. Thus, for s_1 and s_2 having the same sense,

$$g(ks_1, ks_2)=\cos k(s_1-s_2) \tag{16}$$

and for s_1 and s_2 having the opposite sense,

$$g(ks_1, ks_2)=\cos k(s_1+s_2). \tag{17}$$

The function (16) satisfies Eq. (12) with the negative sign, and hence when s_1 and s_2 are of the same sense, from Eq. (15), the contribution Z_1 to the radiation impedance,

$$Z_r=R+jX, \tag{18}$$

becomes

$$Z_1=-j60k\oint_1\oint_2 g(ks_1, ks_2)\times h(ks_1, ks_2)\sin^2\tfrac{1}{2}\theta(s_1, s_2)ds_1ds_2. \tag{19}$$

Similarly, the function (17) satisfies Eq. (12) with the positive sign, and hence when s_1 and s_2 are of the opposite sense, from Eq. (14), the contribution Z_2 to the radiation impedance becomes

$$Z_2=j60k\oint_1\oint_2 g(ks_1, ks_2)\times h(ks_1, ks_2)\cos^2\tfrac{1}{2}\theta(s_1, s_2)ds_1ds_2. \tag{20}$$

Substituting $\theta=0$ into Eq. (19) and $\theta=\pi$ into Eq. (20), it follows that all paths for s_1 and s_2 which are parallel, whether on the same wire or on opposite wires, are eliminated from further consideration in carrying out the integrations for the radiation impedance of a rhombic antenna.

For any pair of wires, the axial and surface paths must be taken along both wires in succession. But due to symmetry, the integrations are the same for the axial path along one wire with the surface path along the other wire as it is if the two paths are interchanged. Hence, the integrations need to be carried out and doubled only for the four cases where the axial path lies either on wire *one* or wire *three* and the surface path lies on either wire *two* or wire *four*.

Thus, from Eqs. (16), (17), (19), and (20),

$$\begin{aligned}Z_r=&\,j120k\int_{-l}^{0}\int_{0}^{l}\cos k(s_1+s_2)e(r_{12})\sin^2\alpha ds_2ds_1\\&-j120k\int_{-l}^{0}\int_{-2l}^{-l}\cos k(s_1-s_2)e(r_{14})\sin^2\alpha ds_2ds_1\\&-j120k\int_{l}^{2l}\int_{0}^{l}\cos k(s_1-s_2)e(r_{23})\sin^2\alpha ds_2ds_1\\&+j120k\int_{l}^{2l}\int_{-2l}^{-l}\cos k(s_1+s_2)e(r_{34})\sin^2\alpha ds_2ds_1,\end{aligned} \tag{21}$$

with

$$\begin{aligned}r_{12}&=(s_2^2+s_1^2+2s_1s_2\cos2\alpha)^{\frac{1}{2}},\\r_{14}&=[(l+s_1)^2+(l+s_2)^2-2(l+s_1)(l+s_2)\cos2\alpha]^{\frac{1}{2}},\\r_{23}&=[(s_1-l)^2+(l-s_2)^2+2(s_1-l)(l-s_2)\cos2\alpha]^{\frac{1}{2}},\\r_{34}&=[(2l-s_1)^2+(2l+s_2)^2-2(2l-s_1)(2l+s_2)\cos2\alpha]^{\frac{1}{2}}.\end{aligned}$$

In the first integral, let $s_1=-x_1$ and $s_2=x_2$; in the second integral, let $s_1=x_1-l$ and $s_2=-(x_4+l)$; in the third integral, let $s_1=x_1+l$ and $s_2=l-x_4$; and in the fourth integral, let $s_1=2l-x_1$ and $s_2=x_2-2l$. Then the radiation impedance becomes

$$Z_r=j240\sin^2\alpha\left[\int_0^l\int_0^l\cos k(x_2-x_1)e(r_{21})dx_2dx_1-\int_0^l\int_0^l\cos k(x_4+x_1)e(r_{41})dx_4dx_1\right] \tag{22}$$

with

$$r_{21}=[x_1^2+x_2^2-2x_1x_2\cos2\alpha]^{\frac{1}{2}}=[(x_1-x_2\cos2\alpha)^2+(x_2\sin2\alpha)^2]^{\frac{1}{2}}, \tag{23}$$

$$r_{41}=[x_1^2+x_4^2+2x_1x_4\cos2\alpha]^{\frac{1}{2}}=[(x_1+x_4\cos2\alpha)^2+(x_4\sin2\alpha)^2]^{\frac{1}{2}}. \tag{24}$$

Upon writing the integrands in the exponential form,

$$Z_r=-j120k\sin^2\alpha(I_1+I_2-I_3-I_4) \tag{25}$$

with

$$I_1=\int_0^l\int_0^l\exp[-jk(x_4+x_1+r_{41})]r_{41}^{-1}dx_4dx_1, \tag{26}$$

$$I_2=\int_0^l\int_0^l\exp[jk(x_4+x_1-r_{41})]r_{41}^{-1}dx_4dx_1, \tag{27}$$

$$I_3=\int_0^l\int_0^l\exp[jk(x_2-x_1-r_{21})]r_{21}^{-1}dx_2dx_1, \tag{28}$$

$$I_4=\int_0^l\int_0^l\exp[-jk(x_2-x_1+r_{21})]r_{21}^{-1}dx_2dx_1. \tag{29}$$

Due to symmetry, x_2 and x_1 may be interchanged in I_4, yielding an equality of I_3 and I_4.

Recalling that the path for x_1 is along the axis of wire *one* whereas the paths for x_2 and x_4 are along the inner peripheries of wires *two* and *four*, respectively, if it is assumed that the axes of the wires are of length l, then the limits for x_2 are actually from $a\cdot\cot\alpha$ to $l-a\cdot\tan\alpha$, and those for x_4 are actually from $a\cdot\tan\alpha$ to $l-a\cdot\cot\alpha$. Hence, finally (Fig. 3)

$$Z_r=-j120k\sin^2\alpha(I_1+I_2-2I_3) \tag{30}$$

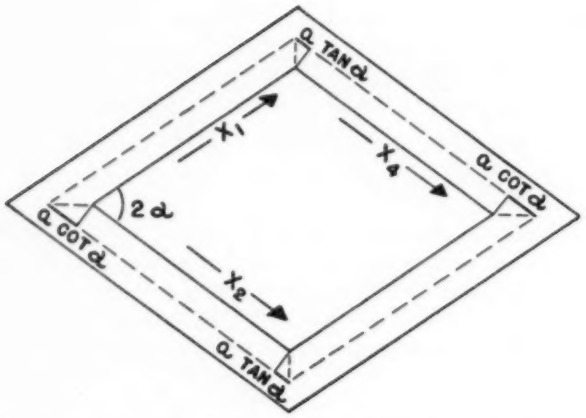


FIG. 3. Cross section of rhombus, showing limits of integration.

with

$$I_1=\int_0^l\int_{a\tan\alpha}^{l-a\cot\alpha}\exp[-jk(x_4+x_1+r_{41})]r_{41}^{-1}dx_4dx_1, \quad (31)$$

$$I_2=\int_0^l\int_{a\tan\alpha}^{l-a\cot\alpha}\exp[jk(x_4+x_1-r_{41})]r_{41}^{-1}dx_4dx_1, \quad (32)$$

$$I_3=\int_0^l\int_{a\cot\alpha}^{l-a\tan\alpha}\exp[-jk(x_2-x_1+r_{21})]r_{21}^{-1}dx_2dx_1. \quad (33)$$

The integrations for I_1, I_2, and I_3 may be carried out by transformations similar to those suggested by F. H. Murray.[2] However, they should be evaluated directly and not by transforming to the exact integral given by Murray, since there is a likelihood of unsuspectingly integrating through a singularity. In this way, the following formula is obtained for the radiation impedance of a rhombic antenna in free space.‡

$$\begin{aligned}Z_r/120=&\,2C+2\ln(2kl\sin^2\alpha)+2Ci2kl\\&-2Ci(2kl\sin\alpha)-Ci[2kl(1+\cos\alpha)]\\&-Ci[2kl(1-\cos\alpha)]\\&+\cos(2kl\sin^2\alpha)\{Ci[2kl\cos\alpha(1+\cos\alpha)]\\&+Ci[2kl\cos\alpha(1-\cos\alpha)]\\&+Ci[2kl\sin\alpha(1+\sin\alpha)]\\&+Ci[2kl\sin\alpha(1-\sin\alpha)]\\&-2Ci[2kl\cos^2\alpha]-2Ci[2kl\sin^2\alpha]\}\\&-\sin(2kl\sin^2\alpha)\{Si[2kl\cos\alpha(1+\cos\alpha)]\\&-Si[2kl\cos\alpha(1-\cos\alpha)]\\&-Si[2kl\sin\alpha(1+\sin\alpha)]\\&+Si[2kl\sin\alpha(1-\sin\alpha)]\\&-2Si[2kl\cos^2\alpha]+2Si[2kl\sin^2\alpha]\}\\&+j[Si[2kl(1+\cos\alpha)]-Si[2kl(1-\cos\alpha)]\\&+2Si[2kl\sin\alpha]-2Si2kl\\&-\cos(2kl\sin^2\alpha)\{Si[2kl\cos\alpha(1+\cos\alpha)]\\&+Si[2kl\cos\alpha(1-\cos\alpha)]\\&+Si[2kl\sin\alpha(1+\sin\alpha)]\\&+Si[2kl\sin\alpha(1-\sin\alpha)]\\&-2Si[2kl\cos^2\alpha]-2Si[2kl\sin^2\alpha]\}\\&-\sin(2kl\sin^2\alpha)\{Ci[2kl\cos\alpha(1+\cos\alpha)]\\&-Ci[2kl\cos\alpha(1-\cos\alpha)]\\&-Ci[2kl\sin\alpha(1+\sin\alpha)]\\&+Ci[2kl\sin\alpha(1-\sin\alpha)]\\&-2Ci[2kl\cos^2\alpha]+2Ci[2kl\sin^2\alpha]\}]. \quad (34)\end{aligned}$$

The resistive component of formula (34) checks with the radiation resistance of a rhombic antenna as given by Leonard Lewin[3] in a discussion of a paper by Donald Foster.[4] From the discussion, it may be inferred that Lewin derived the formula for the radiation resistance by the solid angle Poynting vector method.

THE TERMINATED VEE ANTENNA

It is interesting to note in passing that the free space radiation impedance of the terminated Vee antenna may be written from Eq. (7) as

$$Z_r=60f_0^2\left[\frac{\sin(2kl\sin\alpha)}{2kl\sin\alpha}-1\right]+j120k\sin^2\alpha I_3$$
$$+j120\int_0^l[e(r_{2d})-e(r_{2e})]\sin k(l-s)ds \quad (35)$$

with

$$r_{2d}=[a^2+(l-s)^2]^{\frac{1}{2}}\approx l-s,$$
$$r_{2e}=[l^2+s^2-2ls\cos2\alpha]^{\frac{1}{2}}=[(s-l\cos2\alpha)^2+(l\sin2\alpha)^2]^{\frac{1}{2}}$$

in which each leg of the Vee is of length l, and the vertex angle is 2α.

Using the modified cosine integral function

$$Cin(x)=\int_0^x\frac{1-\cos u}{u}du$$

the radiation impedance of the Vee antenna in free space becomes

$$Z_r/60=f_0^2\left[\frac{\sin(2kl\sin\alpha)}{2kl\sin\alpha}-1\right]$$
$$+2Cin(2kl\sin\alpha)+j2Si(2kl\sin\alpha). \quad (36)$$

In Eq. (36) the end capacitances are assumed to be shunted by the generator and terminal impedance, respectively, and hence have been discarded.

[2] F. H. Murray, Proc. Inst. Radio Engrs. **21**, 1, 154 (1933).

‡ $C=0.5772$, $Si(x)=\int_0^x\frac{\sin u}{u}du$, $Ci(x)=-\int_x^\infty\frac{\cos u}{u}du$.

[3] Leonard Lewin, Proc. Inst. Radio Engrs. **29**, 9, 523 (1941).

[4] Donald Foster, Proc. Inst. Radio Engrs. **25**, 10, 1327 (1937).

The Effect of a Soller Slit on the Diffraction of X-Rays by Deformed Crystals *

B. D. CULLITY AND CARL A. JULIEN †
Department of Metallurgy, University of Notre Dame, Notre Dame, Indiana
(Received October 21, 1952)

In the application of the Guinier-Tennevin focusing Laue method to the study of crystal distortion, it is found that the use of a Soller slit in the incident beam produces striations in reflections from certain crystals. The diffraction geometry of a beam so modified is examined in detail, and it is shown that the striations are due to torsion of the crystal. The relation between the amount of torsion and the slope and spacing of the striations is treated quantitatively. There results a useful extension of the focusing Laue method by which not only the bending, but also the torsion, of a crystal may be measured.

INTRODUCTION

WITHIN a year of the discovery of x-ray diffraction, Bragg[1] noticed a particular feature of the diffracted beams formed in the transmission Laue method: they were convergent in one plane, if the incident beam was divergent in the same plane and the crystal was undistorted. On this basis, Guinier and Tennevin[2,3] devised a focusing Laue method for the detection and measurement of crystal distortion, a method of great sensitivity in disclosing angular disorientation of reflecting lattice planes.

The geometry of the Guinier-Tennevin method is illustrated in Fig. 1. X-rays diverge from the source S, which should be as small as possible, are diffracted by the lattice planes shown and come to a focus at I. The relation between the source-crystal distance SO, measured along the median ray of the beam, and the crystal-focus distance OI, is given by

$$OI = SO\frac{\cos(\alpha+2\phi+\epsilon)}{\cos(\alpha+\epsilon)}, \qquad (1)$$

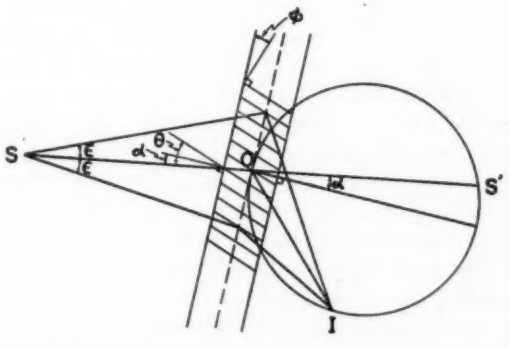


FIG. 1. Focused Laue reflection from a perfect crystal. Not to scale.

where α = angle between the median ray of the incident beam and the normal to the crystal surface, ϕ = angle between the normal to the reflecting lattice planes and the crystal surface, ϵ = angle between median and boundary rays of incident beam. As ϕ changes, the focal point I moves along the circle shown.

If the source S is a fine line, perpendicular to the plane of the drawing, then the focused reflection at I will also be a fine line if the crystal is undistorted. The width of the reflection is approximately equal to the projected width of the source, if ϕ is small and the crystal no thicker than about 0.5 mm. On the other hand, if the crystal is distorted, e.g., by bending about an axis through 0 and normal to the plane of Fig. 1, then the reflection at I becomes broadened by an amount dependent on the disorientation of the reflecting lattice planes. If the distances SO and OI are made rather large, up to 100 cm, disorientation as small as 20 seconds of arc may be detected.[2]

The purpose of the present paper is to describe an extension of the G-T method which allows one to obtain

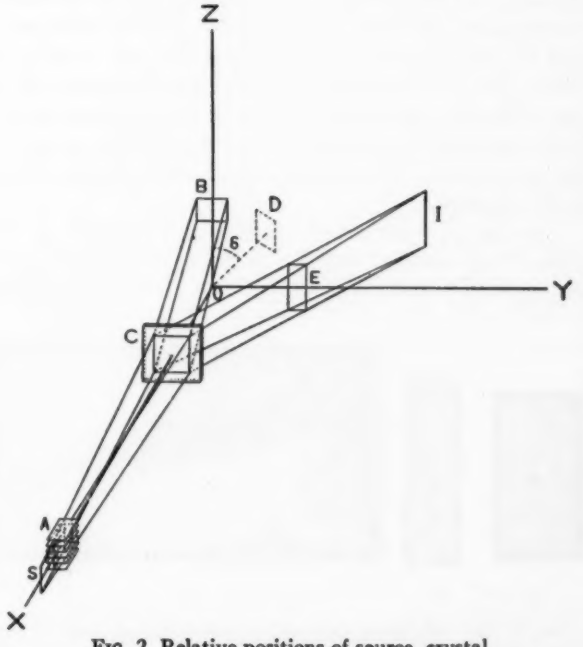


FIG. 2. Relative positions of source, crystal, and various reflections.

* This paper forms part of a dissertation presented by Carl A. Julien to the Engineering Division of the Graduate School of the University of Notre Dame in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

† Now with the Minerals and Metals Advisory Board, National Research Council, Washington, D. C.

[1] W. L. Bragg, Proc. Cambridge Phil. Soc. 17, 43 (1912).

[2] A. Guinier and J. Tennevin, Acta. Cryst. 2, 133 (1949).

[3] A. Guinier and J. Tennevin, *Progerss in Metal Physics* (Interscience Publishers, New York, 1950), vol. 2.

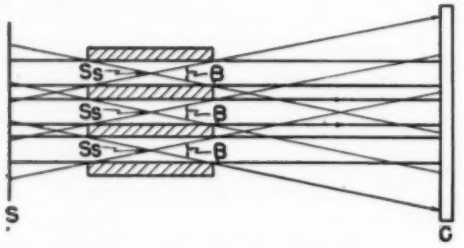


FIG. 3. Vertical section through Soller slit, source (S) and crystal (C). Not to scale.

considerably more information about the nature and extent of the distortion in a crystal.

APPARATUS

Unfiltered copper radiation from a General Electric CA-7 tube was used. This tube has a fine line focus with a projected width, at a target-to-beam angle of 4°, of less than 0.1 mm and a length of about 15 mm.

The relative positions of source and crystal are shown in Fig. 2 in reference to rectangular coordinate axes XYZ. The XY plane and the incident beam, which proceeds in the direction SO, are horizontal; the x-ray source S and the surface of the crystal C are vertical. A Soller slit A is placed near the source. This slit provides for a definite horizontal divergence angle of 3° and contains a set of closely spaced horizontal metal plates designed to reduce the vertical divergence to as low a value as possible. The slit is shown schematically in Fig. 3. The actual slit contains 21 openings, each 1.25 in. long and 0.017 in. high; the maximum angle of vertical divergence β is thus 1.6°. Distinct layers of radiation can be seen on a fluorescent screen held near the slit exit, but these fuse into a single beam a few centimeters away, because of the slight vertical divergence and the fact that the metal plates are only 0.002 inch thick. The source-crystal distance was generally 30–50 cm, while the crystal-film distance varied from about 5 to 50 cm depending on the kind of information sought in the photograph. Exposure times varied from 20 seconds to 20 minutes.

Crystals of quartz and aluminum were used in this work. They were 0.2 to 1.0 mm thick, from 1 to 4 cm in length, and 1 to 4 cm in width. The aluminum crystals were of high purity (99.99 percent Al) and were made by slow solidification of the liquid.

Devices were constructed to produce pure bending or pure torsion in rectangular crystal plates and to permit x-ray examination of the specimen under stress. In each case, the forces were applied by the action of a screw and the strains were calculable from the amount of travel of the screw.

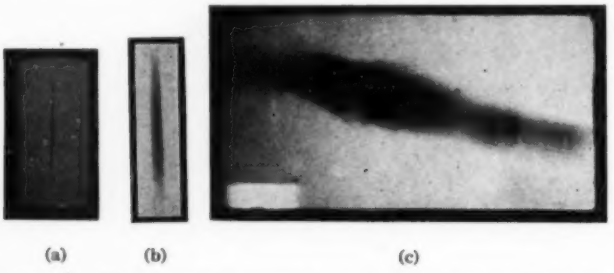


FIG. 4. 90° reflections recorded at the focusing position. (a) Unstrained quartz, (b) quartz bent around vertical axis, (c) inhomogeneously deformed aluminum.

PROCEDURE AND RESULTS

Preliminary

The operation of the G-T focusing Laue method is illustrated in Fig. 2. The crystal C is arranged in the incident beam with its surface vertical and approximately normal to the median ray of the incident beam. A survey photograph is then made by placing the film a few centimeters behind the crystal, i.e., in the YZ plane of Fig. 2. Such a photograph will show a number of large Laue spots, since the film is close to the crystal and the illuminated area of the latter is large (1 to 2 cm²).

If the crystal is correctly oriented for the G-T method, a Laue spot will appear at position E of Fig. 2, i.e., on the equator (Y axis) of the film and just to the right of the spot formed by the transmitted incident beam. If not, the crystal is rotated in its own plane and/or about a vertical axis until a spot appears at E. The reflecting planes involved are then vertical and make small Bragg angles with the incident rays. The position of the spots on such a photograph may be conveniently designated by their azimuthal angle δ relative to a vertical line through the primary beam spot: thus, the reflections B, D, and E are called 0°, δ°, and 90° reflections, respectively.

When the crystal is oriented to produce a spot at E on the large beam Laue photograph, the film is moved back to the focus I, whose position is calculated from Eq. (1), and the G-T reflection recorded. This reflection is considerably behind the YZ plane, with its center in the same horizontal plane (the XY plane) as the center of the line source S. Figure 4(a) shows such a reflection from an undistorted quartz crystal. Its width is no larger than that of the line source, which indicates a very sharp focus and a high degree of perfection in the crystal. Figure 4(b) shows how this reflection is broadened by very slight bending about a vertical axis. Figure 4(c) shows another broad G-T (or 90°) reflection‡ from an aluminum crystal accidentally and inhomogeneously deformed by a small amount. The total disorientation of the reflecting lattice planes in this specimen may be calculated from the total horizontal width of the reflection.

The unexpected feature of the reflection of Fig. 4(c)

‡ A reflection is here called a 90° reflection if it is formed by a reflected beam lying in the horizontal XY plane, whether it is recorded on a film placed at the G-T focusing position or at any other position. Similarly, a 0° reflection is one formed by a beam lying in the vertical XZ plane, whatever the position of the film on which it is recorded.

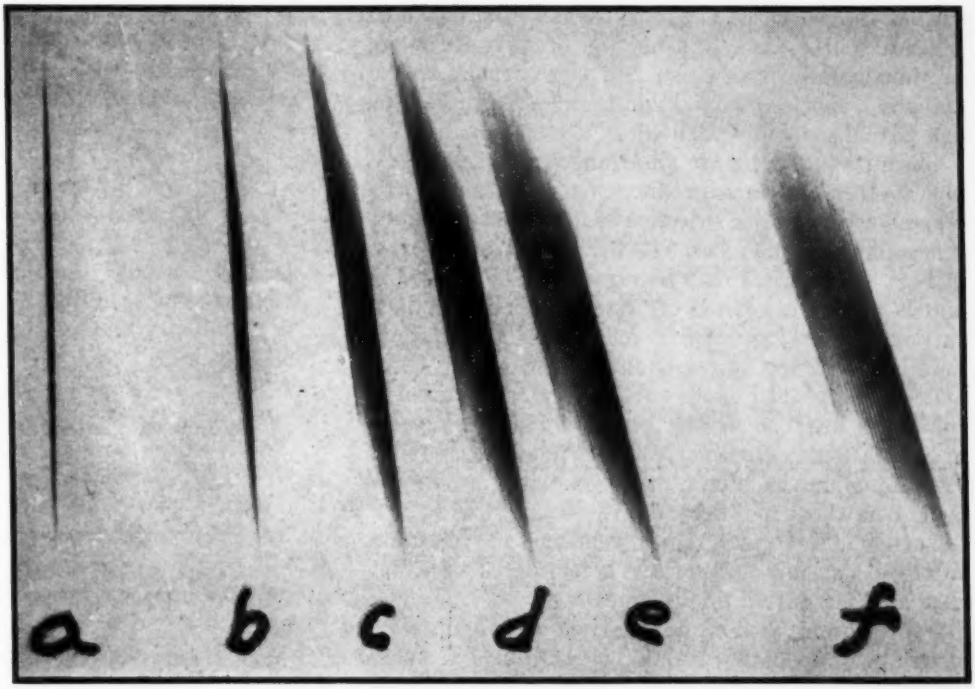


FIG. 5. Focused 90° reflections from quartz, (a) Unstrained. (b) to (f) increasing amounts of torsion about vertical axis. Magnification 3×.

is the wealth of internal detail it contains. What differentiates it from the reflection of Fig. 4(b) is not only the extent of the broadening but also the existence of several sets of fine parallel striations, some straight, some curved, and inclined at various angles to the horizontal. The number and spacing of these striations suggested that they were due in some way to the use of a Soller slit in the incident beam, and this view was easily confirmed by replacing the Soller slit with a conventional one. The reflection remained the same size, but the striations disappeared and were replaced by a more or less uniform blackening of the film.

No ready explanation of these striations could be found. It is true that the beam leaving the Soller slit (Fig. 3) contains horizontal layers of radiation, which show up as strips on a film or fluorescent screen placed near the slit exit, but, as mentioned previously, the vertically divergent rays which are also present quickly overlap these horizontal layers with the result that all the rays merge into a single beam a few centimeters from the slit exit. Yet striations, tilted out of the horizontal, are observed in the *diffracted beam* at distances as large as 70–80 cm from the slit.

EFFECT OF TORSION

An intensive study, theoretical and experimental, was made of these striated reflections, and it was finally established that the striations were caused by torsion of crystal and the presence of a Soller slit in the primary beam. Experimental evidence for this result is given in Fig. 5, obtained from a quartz crystal plate, twisted about its longitudinal axis. The torsion axis was vertical i.e., parallel to the Z axis of Fig. 2. Figure 5 shows the effect of increasing amounts of torsion on the focused 90° reflection. There are six exposures in this photograph, obtained by shifting the film sideways in its own plane after each exposure, and the angle of twist per unit length along the torsion axis varies from zero in Fig. 5(a) to 3.75 min/mm in (f). Resolved striations (very fine, parallel straight lines) become visible in (d) and increase in spacing and inclination to the vertical as the amount of torsion increases in (e) and (f). The striations in reflections from quartz are straighter and more uniform than those in reflections from aluminum [Fig. 4(c)], simply because the former crystal deforms more uniformly than the latter.

The explanation of this phenomenon lies partly in the way in which the incident beam is modified by a Soller slit and partly in the way this modified beam is diffracted by a twisted crystal. In effect, a Soller slit creates a set of virtual x-ray sources. Figure 3 shows that a large number of rays cross at the point S, in the center of each cell, so that the vertically divergent radiation emerging from each cell may be said to issue from a

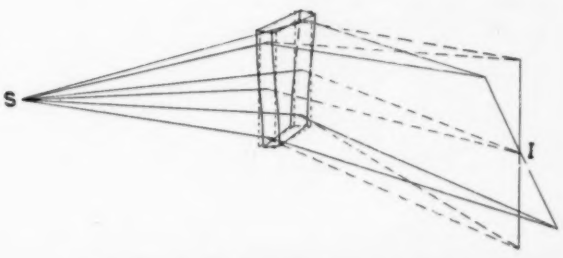


FIG. 6. Focusing of radiation from a point source by a twisted crystal.

virtual source located at S_s rather than from the actual physical source S. At the same time the rays are diverging horizontally from S, so that the virtual source is actually a short horizontal line through S_s, there being one virtual source for each cell of the Soller slit.

Since these sources are all arranged vertically above one another and each produces (from an undistorted crystal) a vertical line reflection at the focusing position, it is apparent that a focused G-T reflection such as those in Figs. 4(a) or 5(a) may be regarded as composed of, and is in fact composed of, a set of vertical colinear "striations" or focused line reflections. There are 21 lines in such a set, corresponding to the 21 cells in the Soller slit, and their centers are vertically separated by the same distance as the centers of adjacent cells, namely 0.48 mm. However, each of these lines is so long, several millimeters in fact, that it overlaps several of its meighbors above and below.

Torsion of the crystal radically changes the reflection. In Fig. 6, a single point source S is considered for simplicity, and the dotted lines delineate the wedge-shaped diffracted beam from an undistorted crystal. The rays of this beam focus in a vertical line through I. When the crystal is twisted, the diffracted beam is also twisted and so is the focused line reflection, as shown by the full lines in Fig. 6, since torsion increases the Bragg angles in the top half of the crystal and decreases them in the bottom half. When there are a number of vertically separated sources, the focused reflection from an undistorted crystal is composed of a set of colinear "striations," as in Fig. 5(a), but torsion of the crystal causes each striation to twist or rotate and shift horizontally, resulting in each striation becoming resolved from its neighbor after a certain amount of torsion has been applied, as in Fig. 5(f).

In Fig. 7, only the median rays of the wedge-shaped incident and diffracted beams are considered. They are shown as planes of radiation vertically diverging from two point sources and as twisted planes after diffraction. S_s is taken as the center of a Soller slit cell, separated from its neighbor by a vertical distance d. If T is the angle of twist per unit length along the torsion axis and a the

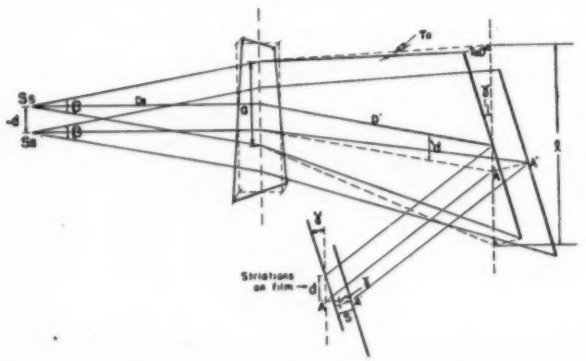


FIG. 7. Separation and tilting of focused reflections by a twisted crystal.

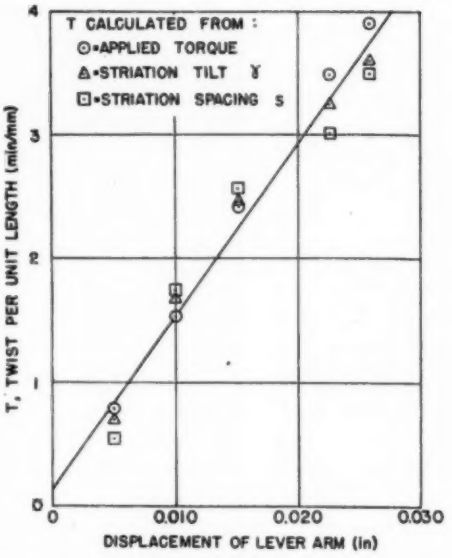


FIG. 8. Comparison of twist per unit length obtained from x-ray and mechanical data.

height of crystal illuminated by each source, then torsion will cause the uppermost and lowermost diffracted beams from each source to be deviated from their original directions by an angle Ta. Let l = height of reflection from untwisted crystal at a crystal-film distance of D' and let γ = angle through which reflection is tilted by torsion. Then

$$\tan\gamma = 2TaD'/l,$$
$$l = a(D_s + D')/D_s.$$

Therefore,

$$T = \frac{(D_s + D')}{2D_sD'} \tan\gamma, \qquad (2)$$

where D_s = distance from crystal to center of Soller slit. This equation shows that the sensitivity of the method in detecting small amounts of torsion, increases with increasing D'. In order to preserve focusing, D_s must also be increased, in accordance with Eq. (1).

T may also be obtained from the striation spacing s, measured at right angles to the striations. If AA' is the horizontal separation of two striation centers, then

$$AA' = 2TdD'$$
$$AA' = d\tan\gamma + (s/\cos\gamma).$$

Therefore,

$$T = \frac{1}{2dD'}\left(d\tan\gamma + \frac{s}{\cos\gamma}\right). \qquad (3)$$

The spacing and tilt of the striations in Fig. 5 were measured and the torque T calculated from Eqs. (2) and (3). In addition, it was possible to calculate torque from the displacement of the end of the lever arm which applied torque to the crystal. The results are compared in Fig. 8, which shows satisfactory agreement between

the torque calculated from the mechanical constants of of the torsion device and that obtained from the x-ray data. Equation (3) does not afford an independent determination of T from s, since it also involves γ, but it does provide a check on the consistency of Eqs. (2) and (3). If desired, an expression for T in terms of s alone can be obtained by substituting Eq. (2) into (3).

X-ray examination of planes inclined to the axis of torsion in quartz led to some interesting results. With the torsion axis inclined at some 15° to the vertical, a focused 90° reflection from a set of vertical lattice planes was obtained with the crystal unstrained. The crystal-focus distance was 31 cm in accordance with Eq. (1). The crystal was then twisted. The reflection was found to become blurred, and the focus was found by trial and error to have moved back to a point 45 cm from the crystal. In this case, therefore, Eq. (1) does not give the correct position of the focus point. Now, it was tacitly assumed above that a horizontal line on the crystal surface remained straight after torsion, and this assumption proved correct when the torsion axis was vertical, since Eq. (1) predicted the observed location of

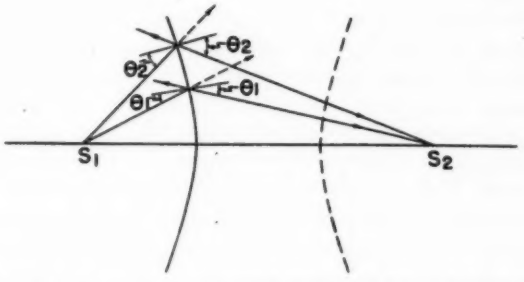
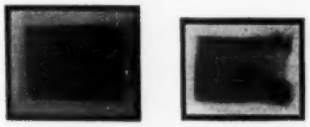


FIG. 9. Focusing by a crystal plate bent into a hyperbola.

the focus point and this location did not change with torsion. However, when the torsion axis is inclined to the vertical, a horizontal line on the crystal surface before torsion must become a curved line concave towards the x-ray source after torsion, since the focus point is moved *away* from the crystal, indicating that the diffracted rays converge less rapidly.

The fact that a focus is obtained at all is surprising, since in general any kind of bending completely disrupts the focusing shown in Fig. 1 and indicates that the concave curve mentioned above must be of some special form. It can be shown analytically that it must be a hyperbola, as illustrated in Fig. 9. The optical analogy here is perfect: light, from one focus S_1, of a hyperbolic mirror will, after reflection, appear to come from the other focus S_2 while x-rays from S_1 will be diffracted to S_2 by lattice planes normal to the surface of a crystal plate bent into the form of a hyperbola. Observation of a sharp focus, therefore, does not of itself prove that a diffracting crystal is undistorted; the criterion for this must be sharp focusing at the distance given by Eq. (1).

0° reflections have interesting properties. These are located on the film vertically above the primary beam



(a) (b)

FIG. 10. Focused 0° reflections. (a) Undistorted quartz, (b) deformed aluminum crystal.

spot and are formed by lattice planes whose normals lie in the XZ plane of Fig. 2. These reflections contain horizontal striations when the crystal is *undeformed*. Figure 10(a) is an example and Fig. 11 shows how the striations are formed; the vertically divergent radiation from each cell of the Soller slit, together with the portion of crystal irradiated thereby, constitutes an individual focusing system. The fact that the crystal-focus distance in this case is related by Eq. (1) to D_s, rather than to the crystal-physical source distance, is proof that the virtual x-ray source does lie at the center of the Soller slit.

The only effect of torsion on a focused 0° reflection is merely to incline the striations slightly to the horizontal. However, some kinds of crystal distortion are nicely revealed in such a reflection. The reflection in Fig. 10(b) was obtained from an aluminum crystal: it shows that left side of the crystal is relatively perfect but that the upper and lower parts of the right side are twisted in opposite directions about a horizontal axis.

Focused reflections having δ-values intermediate between 0° and 90° were also studied, by stepwise rotation of an undistorted crystal about the incident beam as axis. As δ increases, the whole reflection tips and becomes narrower, and the striations move closer together until, at $\delta=90°$, the reflection consists of the usual sharp vertical line.

Reflections formed by CuKα characteristic radiation were also observed. In the 0° position, these reflections are also horizontally striated since they are formed by diffraction of only the horizontal rays in the incident beam. (The vertically diverging rays cannot be diffracted since they do not make the correct Bragg angle.) Such diffraction does not involve focusing and the striations can be observed at any crystal-film distance.

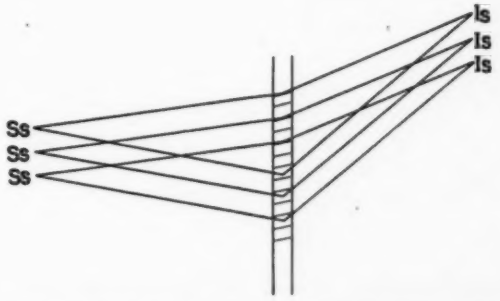


FIG. 11. Vertical section through virtual source S_s, crystal and focused reflection I_s, showing formation of striations in a 0° reflection.

THE EFFECTS OF APERTURES AND EDGES

The striations from twisted quartz in Fig. 5 are sharp and well resolved only when the film is at or near the focusing position, and they become blurred and merge together at other film positions. This is due to the way each wedge of diffracted radiation narrows to a line to form a striation (Fig. 6) and then broadens again beyond the focus point; on a film not sufficiently close to the focus point, adjacent wedges of radiation will overlap to such an extent that separate striations will not be seen. On the other hand, it was observed that the striations in 90° reflections from deformed aluminum crystals were resolvable over a considerable range of film positions.

This fact may be accounted for by assuming that the base of each diffracted wedge of radiation is narrower, horizontally, than the irradiated width of the crystal, thus decreasing the width of the reflected beam at any film position and increasing the distance of no overlap. This condition was realized, with quartz, by placing a shield containing an aperture 1 mm wide in front of a twisted crystal. It was found that the distance over which resolved striations could be detected increased from about 1 cm to about 14 cm. It must be concluded, therefore, that the striations in the reflections from aluminum are caused by rather narrow regions more or less uniformly strained in torsion; these regions of the crystal may be connected, or bounded, by other regions which are bent or bent and twisted. This conclusion was confirmed directly, in one case, by systematically blocking off the incident beam from all but a very small area of an aluminum crystal without finding any change in the striated region of the reflection.

In special cases, striations may be observed in 90° reflections from *undistorted* crystals when a Soller slit is used. These striations have an entirely different origin from the "torsional striations" discussed up to this point and were termed "edge-effect striations," because they can appear whenever an edge of the crystal lies in the incident beam or when an obstructing edge is placed in the incident beam just in front of the crystal. In either case, the effect of the edge is to so modify that part of the incident beam which is diffracted that a series of maxima and minima in intensity are produced over part of the cross section of the diffracted beam. The result is a partly striated Laue reflection. An experienced observer can usually distinguish edge effect striations from torsional striations, but the former can obscure the interpretation of a pattern if one is not alert to the possibility of their existence and takes no steps to prevent their formation.

DISCUSSION

The examples given above show how a Soller slit in the incident beam, in conjunction with a line source of x-rays, adds informative internal detail to Laue reflections. It does so essentially by introducing a set of regularly spaced virtual sources of radiation which irradiate slightly different regions of the crystal.

By the use of quartz as a standard "perfect" crystal, the normal appearance of reflections ranging in position from 0° to 90° was established, both for unstrained crystals and for crystals subjected to controlled elastic strains in bending or torsion. These reflections were then used as standards with which similar reflections from plastically deformed aluminum crystals could be compared in order to assess crystal distortion in the latter. In work to be reported elsewhere, the methods here developed have been used in a study of deformation and recovery in aluminum single crystals. In particular, the use of a Soller slit extends the usefulness of the G-T focusing Laue method by permitting quantitative measurement, not only of bending or fragmentation about one axis, but also of torsion.

ACKNOWLEDGMENTS

The writers are grateful for the support given this research by the University of Notre Dame and by the U. S. Office of Naval Research. They also wish to thank S. S. Hsu for his help in preparing the illustrations.

Current Element Near the Edge of a Conducting Half-Plane

ROGER F. HARRINGTON*
Ohio State University, Columbus, Ohio
(Received October 15, 1952)

The two-dimensional problem of the Maxwellian field from a line source of current adjacent to a conducting half-plane is treated by a transform method of solution. Expressions in the form of contour integrals are given for the current induced in the conductor and for the radiated field.

I. INTRODUCTION

BY a transform method of solution, the field from a line current adjacent to a semi-infinite plane conductor is found. The equation considered is an integral equation similar to the type considered by Wiener and Hopf,[1] and applied to electromagnetic problems by Levine and Schwinger.[2] A description of the general method of solution is given in another paper.[3]

Consider an infinite line source, defined by $x=a$, $z=-b$, in the vicinity of a semi-infinite, perfectly conducting plane, defined by $x=0$, $z<0$. The geometry is illustrated by Fig. 1. The field point is specified by the coordinates x and z, or, alternatively, by the distance r from the origin and the angle θ measured from the positive z axis. The line current is of uniform excitation, taken as unity. From symmetry, the field is everywhere independent of the y coordinate, reducing the problem to a two-dimensional one.

II. DERIVATION OF THE TRANSFORM EQUATION

The field is transverse magnetic to the y axis, and can therefore be represented by

$$H=\nabla\times A,$$
$$j\omega\epsilon E=\nabla\times\nabla\times A=\beta^2 A, \tag{1}$$

where $A=y\psi$, and ψ is a solution to the scalar wave equation. The boundary conditions for the problem are

$$E_y=\beta^2\psi=0 \text{ at } x=0,\ z<0, \tag{2}$$

and the field satisfies the Sommerfeld radiation condition at large distances from the source.

Let ψ be expressed in terms of the incident field from the line source, ψ^i, and a scattered field caused by the presence of the plane conductor, ψ^s, as

$$\psi(x,z)=\psi^i(x,z)+\psi^s(x,z). \tag{3}$$

The incident field is simply the field from a line source radiating into free space, which is

$$\psi^i=\frac{1}{4j}H_0^{(2)}\{\beta[(x-a)^2+(z+b)^2]^{\frac{1}{2}}\}. \tag{4}$$

The scattered field is that produced by the electric currents flowing on the plane conductor. These currents are equal in magnitude to the discontinuity in the tangential magnetic field over the conductor. Since the only component of the tangential magnetic field is H_z the electric currents flow entirely in the y direction. Letting $J(z)$ represent the electric current density, the scattered field can be constructed as the sum of the contributions from differential line elements of current, as

$$\psi^s=\frac{1}{4j}\int_{-\infty}^{0}J(z')H_0^{(2)}\{\beta[x^2+(z-z')^2]^{\frac{1}{2}}\}dz'. \tag{5}$$

The integration extends only over negative values of z' since J is zero for positive z'.

The Laplace transformation is now applied to Eq. (3). A Laplace transform pair is given formally by

$$\bar{F}(\gamma)=\int_{-\infty}^{\infty}F(z)e^{-\gamma z}dz,$$
$$F(z)=\frac{1}{2\pi j}\int_{-j\infty}^{j\infty}\bar{F}(\gamma)e^{\gamma z}d\gamma. \tag{6}$$

The existence of $\bar{F}(\gamma)$ implies certain restrictions on $F(z)$ and on the path of integration in the γ-plane, which may be displaced from the imaginary axis to insure convergence of the integrals. Taking the transform of Eq. (3) gives

$$\Psi(x,\gamma)=\Psi^i(x,\gamma)+\Psi^s(x,\gamma). \tag{7}$$

Both Ψ^i and Ψ^s contain the transform of the Hankel function, which is known.[4] Thus,

$$\Psi^i=\frac{e^{\gamma b}}{4j}\int_{-\infty}^{\infty}H_0^{(2)}\{\beta[(x-a)^2+(z+b)^2]^{\frac{1}{2}}\}e^{-\gamma(z+b)}dz,$$
$$=\frac{e^{\gamma b}\exp[-j|x-a|(\gamma^2+\beta^2)^{\frac{1}{2}}]}{2j(\gamma^2+\beta^2)^{\frac{1}{2}}}, \tag{8}$$

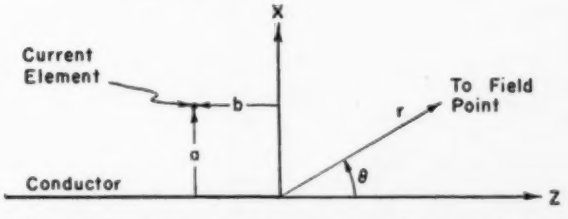


FIG. 1. Current element adjacent to a conducting half-plane.

* Now at Syracuse University, Syracuse, New York.
[1] Paley and Wiener, Am. Math. Soc. Colloq. Pubs. **19** (1934).
[2] Levine and Schwinger, Phys. Rev. **73**, 383–406 (1948).
[3] R. F. Harrington, "On the solution of integral equations of the Wiener-Hopf type." A paper to be published.
[4] Carlson and Heins, Quart. Appl. Math. **4**, 313–330 (1946).

and

$$\Psi^s=\frac{1}{4j}\int_{-\infty}^{0}J(z')e^{-\gamma z'}dz'\times\int_{-\infty}^{\infty}H_0^{(2)}\{\beta[x^2+(z-z')^2]^{\frac{1}{2}}\}e^{-\gamma(z-z')}dz$$

$$=\frac{\exp[-j|x|(\gamma^2+\beta^2)^{\frac{1}{2}}]}{2j(\gamma^2+\beta^2)^{\frac{1}{2}}}\int_{-\infty}^{0}J(z')e^{-\gamma z'}dz',$$

$$=\frac{\exp[-j|x|(\gamma^2+\beta^2)^{\frac{1}{2}}]}{2j(\gamma^2+\beta^2)^{\frac{1}{2}}}\bar{J}(\gamma). \quad (9)$$

Once $\bar{J}(\gamma)$ is determined, substituting Eqs. (8) and (9) into Eq. (7), and inverting, gives the field at any point x, z. Letting $x=0$ in Eq. (7) gives

$$2j(\gamma^2+\beta^2)^{\frac{1}{2}}\Psi(0,\gamma)=e^{\gamma b}\exp[-ja(\gamma^2+\beta^2)^{\frac{1}{2}}]+\bar{J}(\gamma), \quad (10)$$

which is the appropriate transform equation for the problem.

III. TRANSFORM SOLUTION

The regions of regularity of the unknowns must first be determined. From the boundary condition, Eq. (2), it is seen that $\psi(0,z)=0$ for $z<0$, so that

$$\Psi(0,\gamma)=\int_{0}^{\infty}\psi(0,z)e^{-\gamma z}dz. \quad (11)$$

Since ψ is a solution to the two-dimensional wave equation with the source in the finite region, the behavior of ψ for large z is given by

$$\psi(0,z)\underset{z\to\infty}{\longrightarrow}\frac{e^{-j\beta z}}{z^{\frac{1}{2}}}. \quad (12)$$

This implies that $\Psi(0,\gamma)$ is regular in the region $Re(\gamma)>Im(\beta)$. Since $J(z)=0$ for $z>0$, its transform is given by

$$\bar{J}(\gamma)=\int_{-\infty}^{0}J(z)e^{-\gamma z}dz. \quad (13)$$

Now J is equal to the discontinuity in H_z, and is thus a solution to the wave equation. Therefore

$$J(z)\underset{z\to-\infty}{\longrightarrow}\frac{e^{j\beta z}}{(-z)^{\frac{1}{2}}}, \quad (14)$$

which implies that $\bar{J}(\gamma)$ is regular in the region $Re(\gamma)<-Im(\beta)$. With these results, it is seen that all terms of Eqs. (10) are regular in the strip enclosing the imaginary axis, $|Re(\gamma)|<-Im(\beta)$.

It is also necessary to know that the unknowns are well behaved, that is, do not increase exponentially, as $\gamma\to\infty$ in their respective half-planes of regularity. That this condition is met is implied in the assumption that the field transforms exist. In order for the unknowns to be integrable about the z origin, the singularities must be no greater than $1/z$. This insures that both $\bar{J}$ and Ψ do not become infinite faster than the first power of γ as $\gamma\to\infty$.

The function multiplying Ψ in Eq. (10) is now separated into two factors regular in the right and left half-planes, respectively. When this is done, rearranging Eq. (10) gives

$$2j(\gamma+j\beta)^{\frac{1}{2}}\Psi(0,\gamma)-\frac{1}{(\gamma-j\beta)^{\frac{1}{2}}}\bar{J}(\gamma)=\frac{e^{\gamma b}\exp[-ja(\gamma^2+\beta^2)^{\frac{1}{2}}]}{(\gamma-j\beta)^{\frac{1}{2}}}. \quad (15)$$

The first term on the left side of Eq. (15) is regular in the right half-plane and the second term is regular in the left half-plane. Thus, from transform theory, the inverse of the first term of Eq. (15) vanishes for $z<0$, and the inverse of the second term vanishes for $z>0$. The inverse transform of the term on the right of Eq. (15) exists over all z. Therefore, through inversion of Eq. (15), one has

$$\int_{-j\infty}^{j\infty}\frac{\exp[-ja(\gamma^2+\beta^2)^{\frac{1}{2}}]}{(\gamma-j\beta)^{\frac{1}{2}}}e^{\gamma(z+b)}d\gamma=\begin{cases}2j\displaystyle\int_{-j\infty}^{j\infty}(\gamma+j\beta)^{\frac{1}{2}}\Psi(0,\gamma)e^{\gamma z}d\gamma, & z>0,\\ -\displaystyle\int_{-j\infty}^{j\infty}\frac{1}{(\gamma-j\beta)^{\frac{1}{2}}}\bar{J}(\gamma)e^{\gamma z}d\gamma, & z<0.\end{cases} \quad (16)$$

Since the Laplace transformation is unique, the unknowns $\bar{J}$ and Ψ can be found by transforming Eq. (16) over the appropriate ranges. Only $\bar{J}(\gamma)$ is of

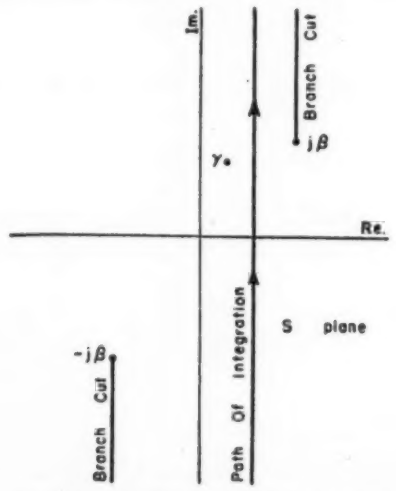


FIG. 2. Path of integration for Eq. (18).

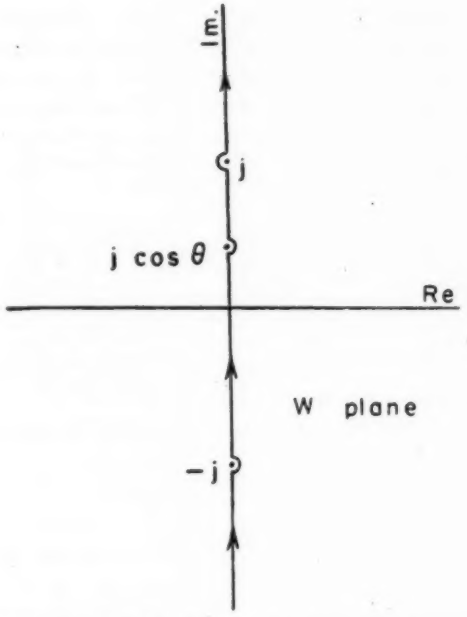


FIG. 3. Path of integration for Eq. (24).

interest now, and is given by

$$J(\gamma)=\frac{-(\gamma-j\beta)^{\frac{1}{2}}}{2\pi j}\int_0^{\infty}\int_{-j\infty}^{j\infty}\frac{\exp[-ja(s^2+\beta^2)^{\frac{1}{2}}]}{(s-j\beta)^{\frac{1}{2}}}\times e^{sb}e^{(s-\gamma)z}dsdz. \quad (17)$$

The complex variable of integration in Eq. (17) has been replaced by s to avoid ambiguity with γ. The unknown $\Psi(0,\gamma)$ is given by a similar expression. Interchanging the order of integrations in Eq. (17) and integrating with respect to z gives

$$J(\gamma)=\frac{-(\gamma-j\beta)^{\frac{1}{2}}}{2\pi j}\int_{-j\infty}^{j\infty}\frac{\exp[-ja(s^2+\beta^2)^{\frac{1}{2}}]}{(s-j\beta)^{\frac{1}{2}}(s-\gamma)}\times e^{bs}ds,\ Re(s)>Re(\gamma). \quad (18)$$

This is the desired solution to the transform equation. By analytic continuation, it holds for all γ. Figure 2 gives the path of integration in the s plane for Eq. (18). There are branch points at $s=\pm j\beta$, and branch cuts are placed from these points to infinity parallel to the imaginary axis. The path of integration lies to the right of the pole $s=\gamma$.

IV. INTERPRETATION OF THE SOLUTION

The solution of part III, substituted into Eqs. (7) to to (9), and inverted, gives the field at any point in space. However, the complexity of the solution is such as to make it difficult to evaluate the inverse transform. Fortunately, the far field can be obtained directly from the transform solution, without recourse to inversion.

Equations (3) to (5) give the field as a function of the distance from a current element,

$$\rho=[(x-x')^2+(z-z')^2]^{\frac{1}{2}}.$$

For large ρ, one can apply the asymptotic form for the Hankel function,

$$H_0^{(2)}(\beta\rho)\underset{\rho\to\infty}{\longrightarrow}(2/\pi\beta\rho)^{\frac{1}{2}}e^{j(\pi/4-\beta\rho)}. \quad (19)$$

For the far field, the source and each differential element of the current sheet see the same angle θ to the field point. Considering the geometry of Fig. 1, for large r,

$$\rho\cong r-x\sin\theta-z\cos\theta. \quad (20)$$

Substituting the above into Eqs. (4) and (5) gives

$$\psi^i\underset{r\to\infty}{\longrightarrow}\frac{1}{4j}(2/\pi\beta r)^{\frac{1}{2}}e^{j(\pi/4-\beta r)}e^{j\beta(a\sin\theta-b\cos\theta)}, \quad (21)$$

$$\psi^s\underset{r\to\infty}{\longrightarrow}\frac{1}{4j}(2/\pi\beta r)^{\frac{1}{2}}e^{j(\pi/4-\beta r)}\int_{-\infty}^{0}J(z)e^{j\beta z\cos\theta}dz. \quad (22)$$

It is noted that the integral in Eq. (22) is simply the transform of $J(z)$ evaluated at $\gamma=-j\beta\cos\theta$. Therefore, by Eq. (3), the far field is given by

$$\psi=K(r)[e^{j\beta(a\sin\theta-b\cos\theta)}+J(-j\beta\cos\theta)], \quad (23)$$

where $K(r)$ is a proportionality factor dependent only on r. In Eq. (18), setting $\gamma=-j\beta\cos\theta$, and letting $w=s/\beta$, gives

$$J(-j\beta\cos\theta)=\frac{(1+\cos\theta)^{\frac{1}{2}}}{2\pi}\times\int_{-j\infty}^{j\infty}\frac{e^{-j\beta a}(1+w^2)^{\frac{1}{2}}e^{\beta bw}}{(1+jw)^{\frac{1}{2}}(\cos\theta-jw)}dw. \quad (24)$$

For β real, the path of integration in the w plane is given

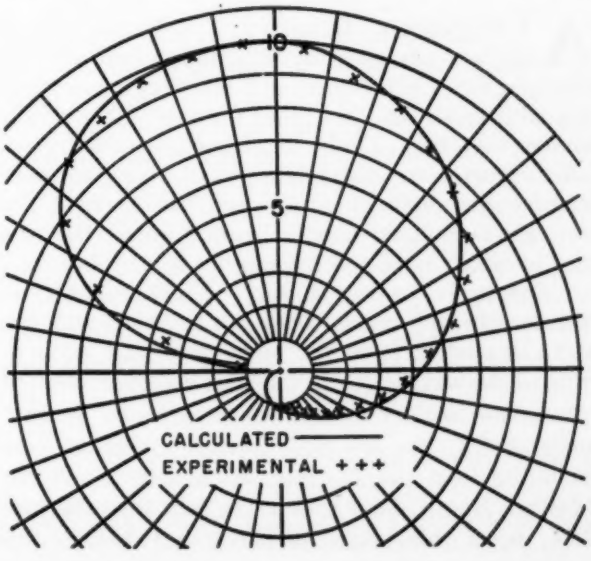


FIG. 4. Field pattern for $\beta a=\lambda/4$, $\beta b=\lambda/4$.

in Fig. 3. The integral in Eq. (24) can be reduced to real integrals, and actual integration carried out numerically. Care must be taken to maintain the proper phase of the radicals.

An interesting check of the solution can be obtained by letting $a=0$, $b>0$ in either the transform solution or the far field solution. This is the situation for which the current element lies on the surface of the conductor. Now the branch point at $\gamma=-j\beta$ is no longer present, and the path of integration can be closed at infinity to the left. Evaluation by the theory of residues shows the field to be everywhere zero, which is the expected result.

The far field pattern given by Eq. (23) is also the pattern of a stub antenna (or finite length line source) in the plane perpendicular to its axis. That this is so can be seen as follows. The far field from an infinite line source is characterized mathematically by the assumption that each differential element of length sees the field point at right angles to its axis. Thus, by superposition, the pattern is also that from a single differential dipole, in the plane perpendicular to its axis.

A pattern has been calculated by numerical integration for the case where the source was a distance $\lambda/4$ from the plane reflector and $\lambda/4$ back from the edge, that is, $\beta a=\lambda/4$, $\beta b=\lambda/4$. Experimental measurements were taken on a model for the same conditions, using a stub antenna for the source. The results are given in Fig. 4. The experimental and theoretical patterns are in good agreement within the experimental accuracy.

JOURNAL OF APPLIED PHYSICS VOLUME 24, NUMBER 5 MAY, 1953

Radiation Field of a Conical Helix

J. S. CHATTERJEE
Institute of Radio Physics and Electronics, Calcutta University, Calcutta, India
(Received November 5, 1952)

It is now well known that a cylindrical helix, when excited at frequencies corresponding to wavelengths comparable to the length of one turn of the helix, can radiate a sharp beam along the axis over a wide frequency range (about one octave). It is shown in the present communication that if the helix be conical instead of cylindrical (the diameter varying along the length of helix), then the axial mode of radiation can be maintained over a much wider band of frequencies. The radiation pattern of a conical helix, 60 cm diameter at the base, tapering linearly to 20 cm at the top in 10 turns within a height of 112 cm (with the "ground" provided by brass disk of 100 cm in diameter) has been studied experimentally. It is found that the axial mode of radiation is maintained from 150 Mc/sec to 450 Mc/sec. By increasing the number of turns, the band width can be considerably increased. Assuming a linear current distribution, theoretical expressions have also been deduced for E_ϕ and E_θ for a conical helix. Some modifications of the simple conical helix, such as may have special applications, are indicated.

I. INTRODUCTION

A CYLINDRICAL helix is a well-known circuit element having many uses. It is widely used, for example, as an inductance; in a traveling wave tube it is used to guide the wave along the axis with a velocity smaller than that of light. In all such applications, the helix diameter is a small fraction of the free space wavelength corresponding to the frequencies concerned. Uses have also been found in recent years of a helix, the diameter of which is of the same order as the free space wavelength. It has been shown by Kraus[1] that for such wavelengths, the circular helix can be an efficient radiator. Depending upon the pitch angle and upon the ratio of the wavelength to the helix diameter, the helix can radiate in three modes. For example, for the pitch angle of 12.6° the so-called "normal" radiation mode (resulting from the T_0 mode of current) is observed when the length of one turn of the helix is smaller than 0.8λ; when this length is between $0.8\lambda-1.3\lambda$, the "axial" radiation mode, due to the T_1 mode of current distribution along the helix, is obtained; and, when the length is more than 1.3λ, energy is radiated in the "conical" mode due also to T_2 mode of current distribution.[2,3] Further, for the axial mode of radiation, a sharp beam is maintained along the axis over the large frequency band of nearly one octave. The three modes[2] of radiation are illustrated in Fig. 1.

It is interesting to inquire how the radiation characteristic of the helix will be altered, if the diameter instead of remaining constant varies along the length of the helix. This problem has been investigated both experimentally and theoretically by the author of the paper for the case of a simple conical helix, the radius of which varies along the length of the axis and the pitch angle remaining constant. It has been found that for such a helix there exists in general all the three modes of radiation as in the case of the cylindrical helix. The normal radiation mode is present when the frequency is

[1] J. D. Kraus, Electronics **20**, No. 4, 109 (1947).

[2] J. D. Kraus and J. C. Williamson, J. Appl. Phys. **19**, 87 (1948).
[3] J. A. Marsh, Proc. Inst. Radio Engrs. **39**, 668 (1951).

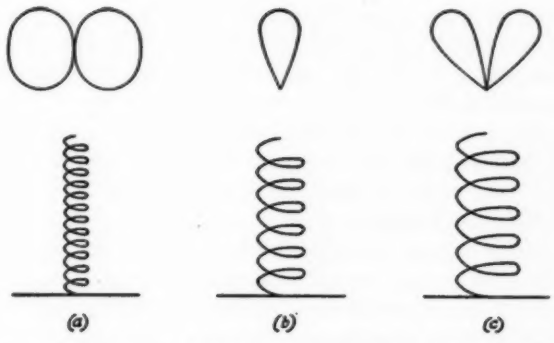


FIG. 1. Three types of radiation pattern of a cylindrical helix: (a) normal mode, (b) axial mode, (c) conical mode.

so low that at no part of the helix the diameter is large enough to satisfy the condition of axial mode of radiation. The axial mode is obtained with higher frequencies. In the transition range of frequencies, the portion of the helix having smaller diameter radiates in the normal mode; of the rest, the part having the larger and the appropriate diameter radiates in the axial mode. At still higher frequencies, the diameter at some portion of the helix becomes large enough (compared to the free space wavelength) for the conical mode of radiation to appear. The relative strengths of these three modes of radiation depend on the frequency and on the position of the point at which energy is introduced for exciting the helical antenna. This is because the feed position determines the current strengths in the different parts of the helix. For example, with the helix used by us for 450 Mc/sec, the radiation is mainly in the axial mode for the "apex feed," the conical mode of radiation just appearing. For the "base feed" the conical mode of radiation is stronger than the axial mode at 400 Mc/sec. Besides other interesting properties, the conical helix can radiate efficiently a beam of small angle over a frequency range which is much greater than that for the corresponding cylindrical helix.

In what follows we shall first describe the results of experimental investigation on the polar diagrams and on the charge distributions along a conical helix of chosen dimensions when it is excited to radiate the different modes. An attempt will then be made to develop a theory of radiation from such a simple conical antenna (Sec. III). Modification of the simple conical helical form to meet the demands of special applications will also be briefly indicated (Sec. IV).

II. EXPERIMENTS WITH A CONICAL HELIX

The form of a simple conical helix can be expressed in cylindrical coordinates (z, a, ϕ) by the equations,

$$a=a_0(1-k\phi) \tag{1}$$

and

$$dz=a\tan\alpha\cdot d\phi, \tag{2}$$

where a is the distance of an element of the helix from the axis.

$$z=\int_0^{\phi} a_0(1-k\phi)\tan\alpha\cdot d\phi=a_0\tan\alpha\left(1-\frac{k\phi}{2}\right)\phi, \tag{3}$$

where α is the inclination of an element of the aerial wire with the horizontal plane and is constant. It is obvious that a and z can be made to vary with ϕ in any way.

The conical helix used by us for studying the charge distribution and the polar diagram of the radiation pattern was made of $\frac{5}{32}$-in. brass rod and had 10 turns. The radius at the base was 30 cm. The radius decreased uniformly to 10 cm at the apex in 10 turns within a height of 112 cm. A brass sheet, 100 cm in diameter, provided the "ground." Measurements were made with the helix oriented in two different ways with respect to the ground. First the base of the cone was close to the ground, and secondly the apex was close to the same. (The distance between the point closest to the ground and the ground was 1 cm in each case.) For each of these two cases polar diagrams were plotted, first, when the feed was at the base, and secondly, when the feed was at the apex. The feed arrangement for each of the cases is illustrated in Fig. 2.

Measurement of the Electric Field

The vertical and horizontal components of the electric field were measured with a small dipole placed at a

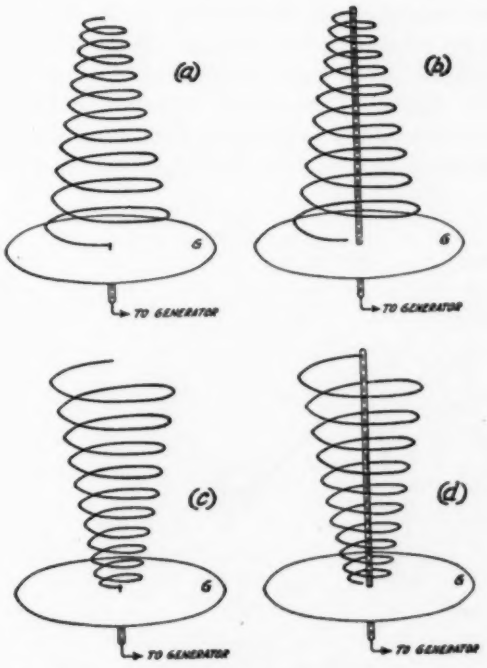


FIG. 2. Illustrating the different methods of feeding energy to the conical helix for different dispositions with respect to "ground." For case (a) the feed is at the base; (b) the feed is at the apex, the feeder passing axially through the helix; (c) feed at the apex; (d) feed is at the base, the feeder passing axially through the helix. Case (b) has the largest useful band width. For case (a) the band width is smaller, but the directivity is stronger.

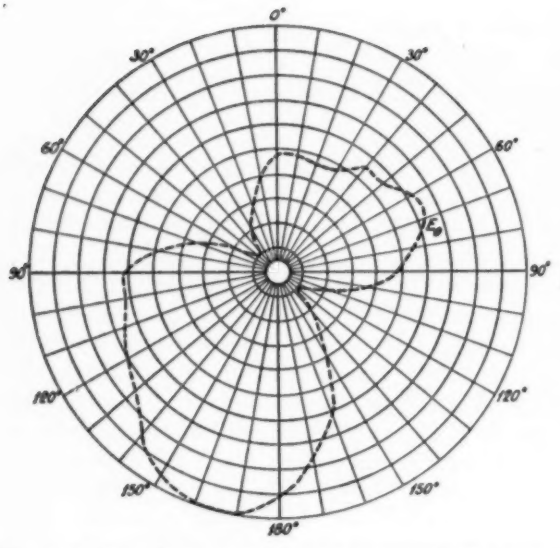


FIG. 3(a). E_θ radiation pattern of the conical helix shown in Fig. 2a. Base is close to the "ground." Energy feed is at the base. Frequency, 120 Mc/sec.

distant point and at the same height as the axis of the aerial. The polar diagrams as obtained for the four cases are shown in Figs. 3–6. The polar diagrams are necessarily different for the four cases because of the different current distributions.

Measurement of the Charge

The current distribution along the antenna can be found by measuring the voltage induced in a balanced loop placed in the axial slot in the aerial as had been done by Marsh.[3] This method, however, is not suitable here as the antenna wire is thin. The charge distribution was measured by coupling capacitatively a small probe to the antenna wire. The coupling was kept small so as not to distort the current distribution. The charge distributions for the top and the bottom feed for the normal and the inverted cone are shown in Figs. 7 and 8. For the case of the top feed, the charge distributions for 100 Mc/sec, 200 Mc/sec, and 300 Mc/sec were measured and for the bottom feed, charge distribution for 200 Mc/sec is shown.

In the case of a circular helix radiating in the axial mode, there are two modes of current propagation, namely, the T_0 and T_1 modes, both in the forward and in the reverse direction. The characteristic of the T_0 mode is that its phase velocity is equal to that of light when the diameter of helix D exceeds 0.5λ. For the T_0 mode, the current is rapidly attenuated and is of negligible amplitude within a few turns. The T_1 mode of current, however, is of constant amplitude along the helix and its phase velocity is smaller than that of light, being a function of pitch angle and of D/λ. The energy reaching the end of a helix of several turns is, thus, that produced by the T_1 mode of current.

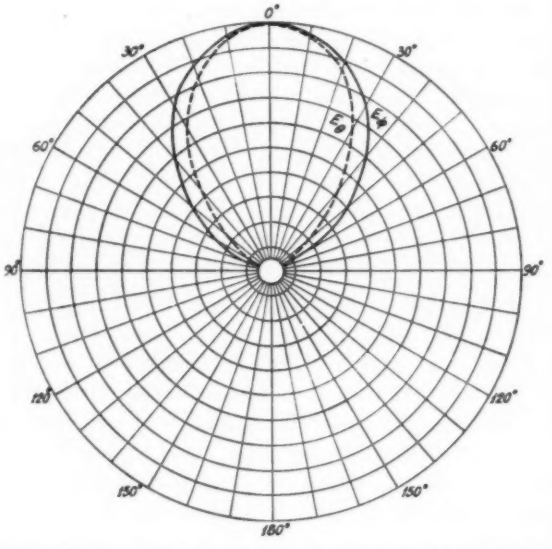


FIG. 3(c). E_ϕ and E_θ radiation patterns of the conical helix shown in Fig. 2a. Base is near the "ground." Feed is at the base. Frequency, 200 Mc/sec.

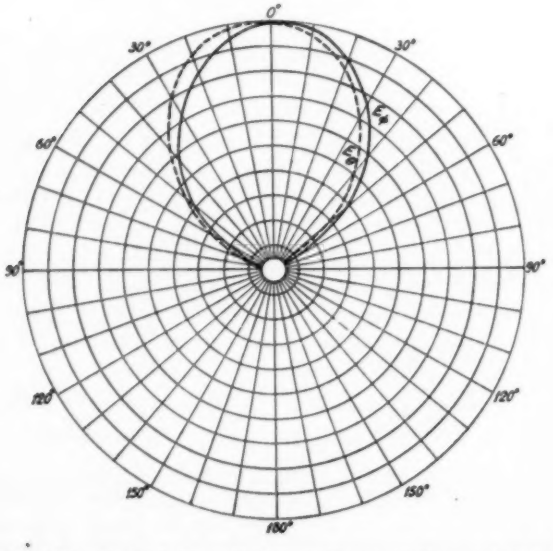


FIG. 3(b). E_ϕ and E_θ radiation patterns of the conical helix shown in Fig. 2a. Base is close to the "ground." Energy feed is at the base. Frequency, 150 Mc/sec.

The current distribution along the length of the conical antenna is, however, much more complicated than that of the cylindrical helix. This is not only because there are several waves traveling with different propagation constants in both directions, but also because the propagation constants of the various modes of current cannot be assumed to be constant along the length of the aerial. In the absence of any theoretical analysis, the following qualitative information regarding the current distributions may be inferred from the experimentally obtained charge distribution. (a) From Fig. 7(a) it is seen that at 100 Mc/sec the T_0 mode is predominant, and pure standing waves are obtained. The propagation velocity is that of light. (b) Figure 7(b)

shows that at 200 Mc/sec the T_1 mode is excited, and the T_0 mode is attenuated. It is also seen that the current is almost zero at the base of the helix. This means that only progressive waves of current exist, and the charge and the current distributions are, therefore, identical. Further, the T_1 mode of current now grows in amplitude along the length of the aerial and is then attenuated. This is to be expected from the results of measurements by Marsh[3] on the circular helix, which indicates that the T_1 mode of current cannot be excited for all the values of D/λ. Since D varies continuously in the case of conical helix, the T_1 mode of current grows and after reaching a maximum amplitude is attenuated. It is also seen that the maximum amplitude of the current is obtained when $D=0.21\lambda$. (c) Figure 7(c) shows the charge distribution for 300 Mc/s. It is of the same nature as Fig. 7(b), except that the maximum amplitude of the T_1 mode is now at the feed point of the helix when $D=0.2\lambda$.

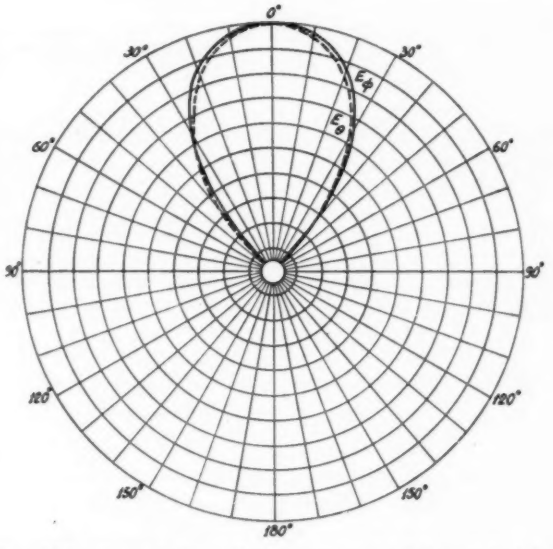


FIG. 3(d). E_ϕ and E_θ radiation patterns of the conical helix shown in Fig. 2a. Base is near the "ground." Feed is at the base. Frequency, 300 Mc/sec.

One important feature of the current distribution is to be noted in each of the cases illustrated. The current distribution is almost zero in the region farther from the feed point, i.e., near the base, for all frequencies higher than that at which the axial mode of radiation appears. This means that the reflection coefficient is zero near the base, or, in other words, the antenna has a constant input impedance over this entire band of higher frequencies.

III. THEORETICAL ANALYSIS

We shall calculate the electric field for the T_1 mode of current distribution. This is because this mode causes the axial radiation which is of chief interest. For the circular helix the electric field at a distant point has been calculated amongst others[2,4] by Kornhauser,[5] by determining the vector potential for the whole of the helix assuming a certain current distribution. Such a procedure is not suitable for the conical helix, because it leads to nonintegrable terms. Instead, the electric fields are calculated for the separate turns, and the total electric field is then obtained by combining them. This, of course, is possible if the distribution of phase velocity is known along the aerial.

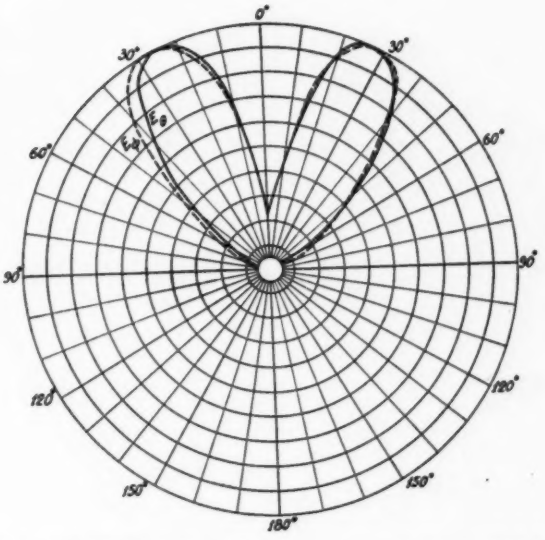


FIG. 3(e). E_ϕ and E_θ radiation patterns of the conical helix shown in Fig. 2a. Base is near the "ground." Feed is at the base. Frequency, 400 Mc/sec.

The electric field E for the far zone is given by

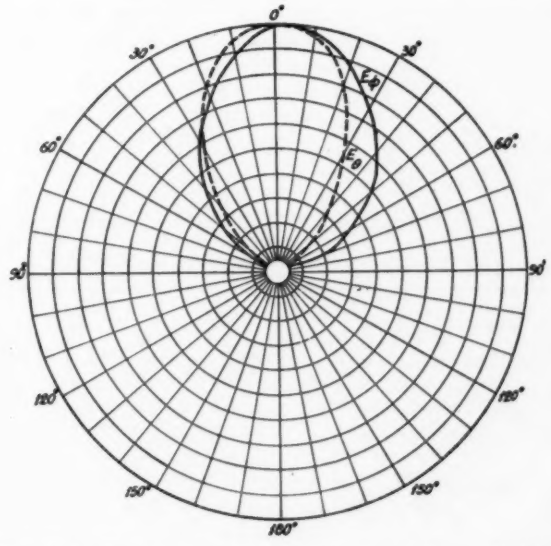


FIG. 4(a). E_ϕ and E_θ radiation patterns of the conical helix shown in Fig. 2b. Base is near the "ground." Feed is at the apex. Frequency, 150 Mc/sec.

[4] H. Lottrup Knudsen, Trans. Danish Acad. Tech. Sci., No. 8, pp. 55 (1950).
[5] E. T. Kornhauser, J. Appl. Phys. 22, 887 (1951).

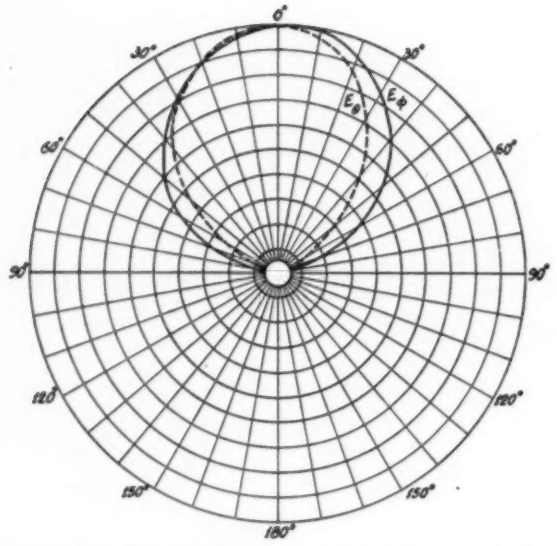


FIG. 4(b). E_ϕ and E_θ radiation patterns of the conical helix shown in Fig. 2b. Base is near the "ground." Feed is at the apex. Frequency, 200 Mc/sec.

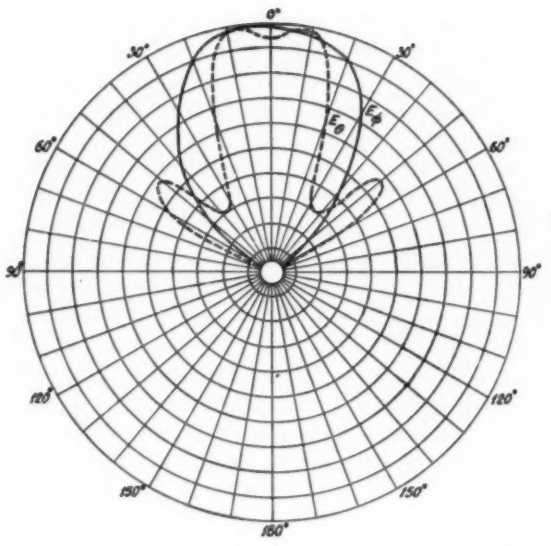


FIG. 4(d). E_ϕ and E_θ radiation patterns of the conical helix shown in Fig. 2b. Base is near the "ground." Feed is at the apex. Frequency, 450 Mc/sec.

(King[6])

$$E=\frac{j\omega\mu_0}{4\pi R}\int\{\mathbf{R}_0\times(\mathbf{R}_0\times\mathbf{S})\}I\cdot\exp(-j\beta_0 R)\cdot ds, \quad (4)$$

where ω is the angular frequency, μ_0 is the absolute permeability of space. R is the distance between the point at which the electric field is calculated and ds, the element of the aerial wire carrying the current I. $\mathbf{R}_0$ is the unit radial vector in a spherical coordinate system (r, θ, ϕ) with the aerial at the origin. $\mathbf{S}$ is the unit vector along the antenna and thus varies in direction along the aerial.

The vector product

$$\{\mathbf{R}_0\times(\mathbf{R}_0\times\mathbf{S})\}=S_\theta\cdot\boldsymbol{\theta}-S_\phi\cdot\boldsymbol{\phi}, \quad (5)$$

where S_θ and S_ϕ are the direction cosines between $\mathbf{S}$ and the θ, ϕ spherical coordinate axes, and are given by

$$S_\theta=-\cos\alpha\cdot\cos\theta\left\{\sin\phi-\frac{k_m\cos\phi}{1-k_m\phi}\right\}-\sin\alpha\cdot\sin\theta,$$

$$S_\phi=\cos\alpha\left\{\cos\phi-\frac{k_m\sin\phi}{1-k_m\phi}\right\}. \quad (6)$$

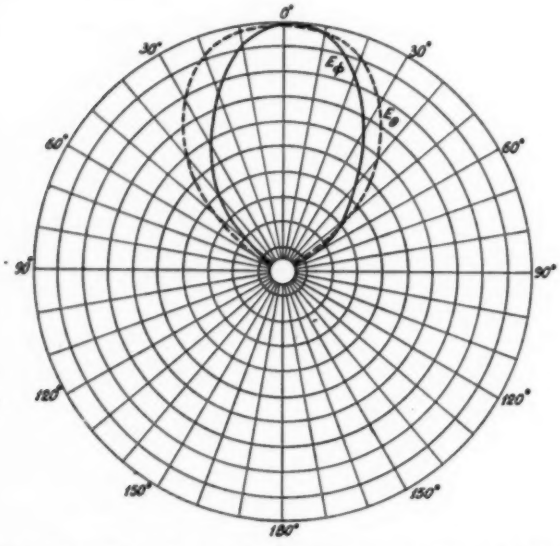


FIG. 4(c). E_ϕ and E_θ radiation patterns of the conical helix shown in Fig. 2b. Base is near the "ground." Feed is at the apex. Frequency, 300 Mc/sec.

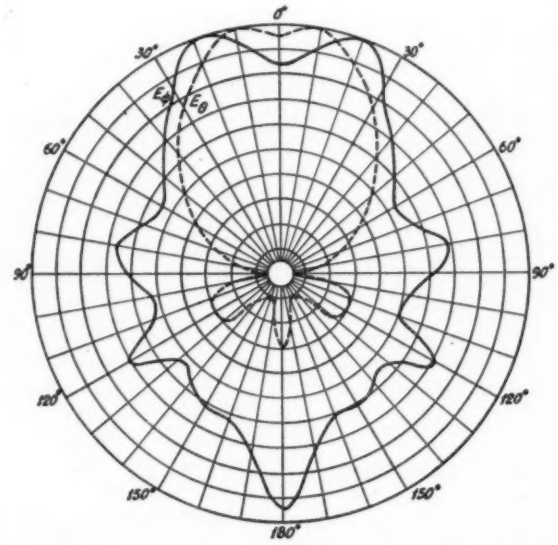


FIG. 5(a). E_ϕ and E_θ radiation patterns of the conical helix with the apex close to the "ground." Feed is at the base. Frequency, 150 Mc/sec.

[6] R. W. P. King, *Electromagnetic Engineering* (McGraw-Hill Book Company, Inc., New York, 1945), Vol. 1, pp. 271.

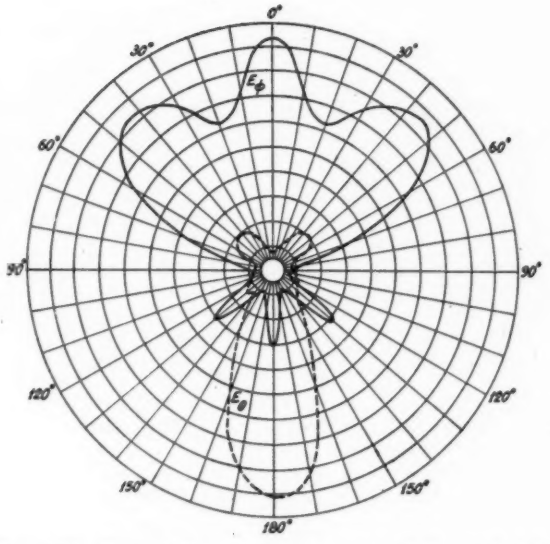


FIG. 5(b). E_ϕ and E_θ radiation patterns of the conical helix with the apex close to the "ground." Feed is at the base. Frequency, 200 Mc/sec.

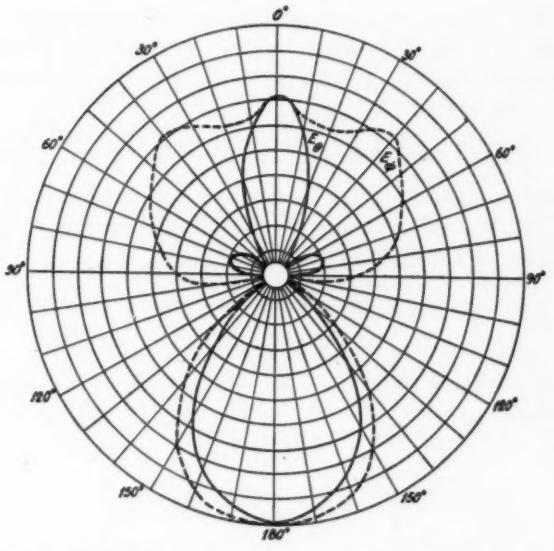


FIG. 6(a). E_ϕ and E_θ radiation patterns of the conical helix. Apex is close to the "ground" and feed is at the apex. Frequency, 150 Mc/sec.

The ϕ-axis of the spherical coordinate system is taken in the same plane as that of the cylindrical coordinate systems.

From the helix geometry

$$s = a_m \sec\alpha \cdot \phi(1 - k_m\phi/2), \tag{7}$$

$$ds = a_m \sec\alpha \cdot (1 - k_m\phi) \cdot d\phi. \tag{8}$$

Since R is large compared to the dimensions of the helix, it is replaced except in the phase term by R_0, the distance between the origin of the coordinate system and the distant point at which electric field is to be calculated. This approximation cannot obviously be made for the phase term. The value of R for the mth loop is given by (Fig. 9)

$$\begin{aligned} R_m &= R_0 - (z_m + z)\cos\theta - a_m(1 - k_m\phi)\sin\theta \cdot \cos\phi \\ &= R_0 - z_m\cos\theta - a_m \tan\alpha \cdot \phi \cdot \cos\theta \\ &\quad + a_m \tan\alpha \cdot \cos\theta \cdot k_m\phi^2/2 - a_m \sin\theta \cdot \cos\phi \\ &\quad + a_m k_m \sin\theta \cdot \phi \cdot \cos\phi, \end{aligned} \tag{9}$$

where z_m and a_m are values of z and a, respectively, at the center of mth loop. k_m is the value of k for the mth loop.

Since the T_1 mode of current attenuates slowly, it is assumed for simplicity of calculation that the current remains constant over one turn and varies from turn

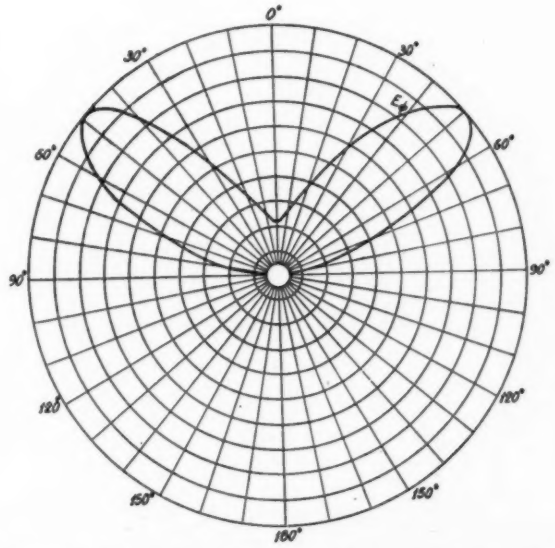


FIG. 5(c). E_ϕ radiation pattern of the conical helix with the apex close to the "ground." Feed is at the base. Frequency, 300 Mc/sec.

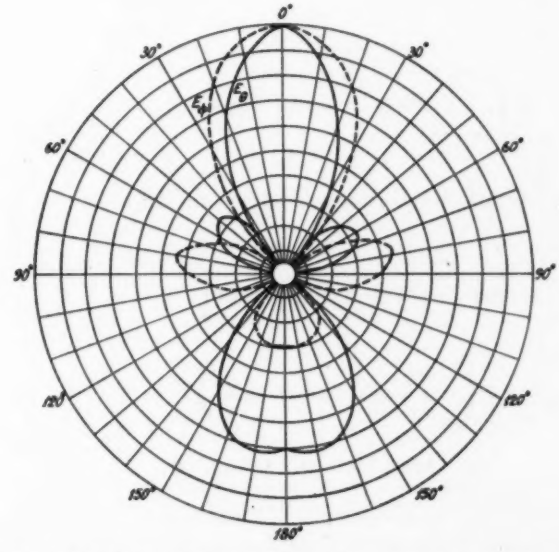


FIG. 6(b). E_ϕ and E_θ radiation patterns of the conical helix. Apex is close to the "ground" and feed is at the apex. Frequency, 200 Mc/sec.

to turn by appropriate amounts. A linear variation of phase velocity is assumed. Thus the current I in the mth loop, referred to the phase of the current in the first loop is given by

$$I = I_m \exp(-jv_m)\cdot\exp(-j\omega s/v), \tag{10}$$

where v_m = phase difference between the current at the centers of first loop and mth loop, and v is the phase velocity and is given by

$$\begin{aligned} v &= v_m(1-h_m\phi) \\ &\doteq v_m/(1+h_m\phi), \end{aligned} \tag{11}$$

since the maximum value of ϕ is π and h_m is small. Substituting Eqs. (6)–(11) in Eq. (5), E_ϕ and E_θ are obtained. Thus,

$$\begin{aligned} E_\phi = -\frac{j\omega\mu_0}{4\pi R_0}\sum_{m=1}^{N}\int_{-\pi}^{+\pi}\Bigg[&\cos\alpha\left\{\cos\phi-\frac{k_m\sin\phi}{1-k_m\phi}\right\} \\ &\times\exp\left(-\overset{\omega}{j}R_m\right)\cdot I_m\cdot\exp(-jv_m) \\ &\times\exp\left\{-j\frac{\omega a_m\sec\alpha}{v_m}\left(\phi-\frac{k_m\phi^2}{2}\right)(1+h_m\phi^2)\right\} \\ &\times a_m\sec\alpha(1-k_m\phi)\Bigg]d\phi, \end{aligned} \tag{12}$$

where N is the number of turns.

Substituting the value of R_m from Eq. (9),

$$\begin{aligned} E_\phi = -\frac{j\omega\mu_0\exp(-j\omega R_0/c)}{4\pi R_0}&\sum_{m=1}^{m=N} a_m I_m \\ &\times\exp\left(\overset{\omega}{j}z_m\cos\theta - jv_m\right) \\ &\times\int_{-\pi}^{+\pi}\Bigg[\{\cos\phi(1-k_m\phi)-k_m\sin\phi\} \\ &\times\exp\left(\overset{\omega}{j}a_m\tan\alpha\cdot\cos\theta\cdot\phi-\overset{\omega}{j}a_m\tan\alpha\right. \\ &\left.\times\cos\theta\cdot\frac{k_m}{2}\phi^2+\overset{\omega}{j}a_m\sin\theta\cdot\cos\phi\right) \\ &\times\exp\left(-\overset{\omega}{j}a_m k_m\sin\theta\cdot\phi\cdot\cos\phi\right) \\ &\times\exp\left\{-\frac{j\omega a_m\sec\alpha\cdot\phi}{v_m}-\frac{j\omega a_m\sec\alpha}{v_m}\right. \\ &\left.\times(h_m-k_m/2)\phi^2+\frac{j\omega a_m h_m k_m}{2v_m}\cdot\phi^3\right\}\Bigg]d\phi. \end{aligned} \tag{13}$$

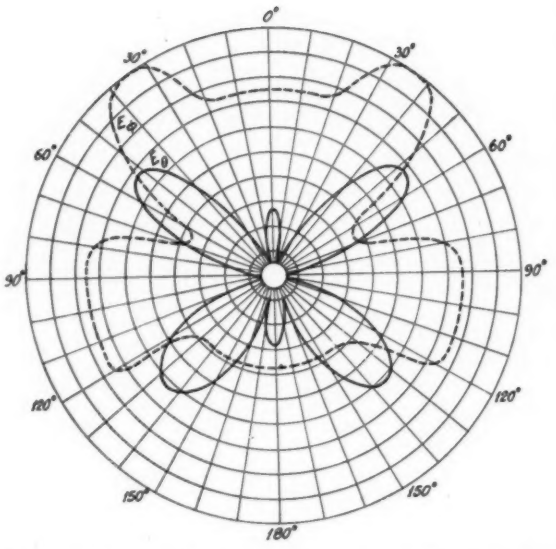


FIG. 6(c). E_ϕ and E_θ radiation patterns of the conical helix. Apex is close to the "ground" and feed is at the apex. Frequency, 300 Mc/sec.

Or, with the abbreviations,

$$E_0 = -\frac{j\omega\mu_0\exp(-j\omega R_0/c)}{4\pi R_0},$$

$$n_{\theta m} = \overset{\omega}{-}z_m\cdot\cos\theta,$$

$$\psi_{\theta m} = n_{\theta m} - v_m,$$

$$\mathfrak{C}_{\theta m} = \overset{\omega}{-}a_m\tan\alpha\cdot\cos\theta,$$

$$d_{\theta m} = \overset{\omega}{-}a_m\sin\theta,$$

$$g_m = \overset{\omega}{-}a_m\sec\alpha. \tag{14}$$

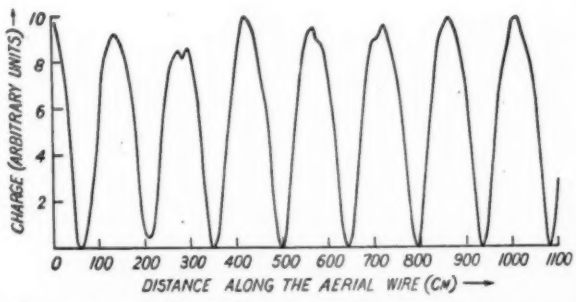


FIG. 7(a). Charge distribution along the length of the conical helix shown in Fig. 2. Base is close to the "ground" and feed is at the apex. Excitation frequency, 100 Mc/sec.

Equation (13) can be written as

$$E_\phi = E_0 \sum_{m=1}^{N} a_m I_m \exp(j\psi_{\theta m}) \int_{-\pi}^{+\pi} \times \Big[\{\cos\phi(1-k_m\phi) - k_m \sin\phi\} \exp j(\mathfrak{C}_{\theta m} - g_m)\phi \times \exp -j\{g_m(h_m - k_m/2) + \mathfrak{C}_{\theta m} \cdot k_m/2\}\phi^2 \times \exp\left(-jg_m h_m \frac{k_m}{2} \cdot \phi^3\right) \times \exp(jd_\theta \cos\phi) \cdot \exp(-jd_\theta k_m \phi \cdot \cos\phi) \Big] d\phi. \quad (15)$$

The right-hand side of the equation cannot in general be integrated. In the particular case where k_m is small, which is of importance, the expression can be simplified. Since the maximum value of ϕ is π, and $g_m h_m k_m \pi^3/2 \ll 1$, we can put,

$$\exp(-jg_m h_m(\tfrac{1}{2}k_m) \cdot \phi^3) \doteqdot 1,$$

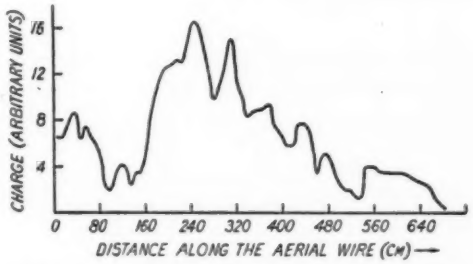


FIG. 7(b). Charge distribution along the length of the conical helix shown in Fig. 2. Base is close to the "ground" and feed is at the apex. Excitation frequency, 200 Mc/sec.

and

$$\exp(-jd_\theta k_m \phi \cos\phi) \doteqdot 1 - jd_\theta k_m \phi \cos\phi \doteqdot 1.$$

Writing,

$$g_m(h_m - k_m/2) + \mathfrak{C}_{\theta m} k_m/2 = H \quad (16)$$

and using the Fourier expansion for $\exp(jd_\theta \cos\phi)$, i.e.,

$$\exp(jd_\theta \cos\phi) = \sum_{n=-\infty}^{n=+\infty} j^n J_n(d_\theta) \exp(jn\phi), \quad (17)$$

where $J_n(d_\theta)$ is the nth order Bessel function for the argument d_θ. Eq. (15) can be written as

$$E_\phi = E_0 \sum_{m=1}^{N} a_m I_m \exp(j\psi_{\theta m}) \int_{-\pi}^{+\pi} \times \Big[\{\cos\phi(1-k_m\phi) - k_m \sin\phi\} \exp\{j(v_{\theta m} - g_m)\phi\} \times (1 - jH\phi^2) \times \sum_{n=-\infty}^{n=+\infty} j^n J_n(d_\theta) \exp(jn\phi) \Big] d\phi.$$

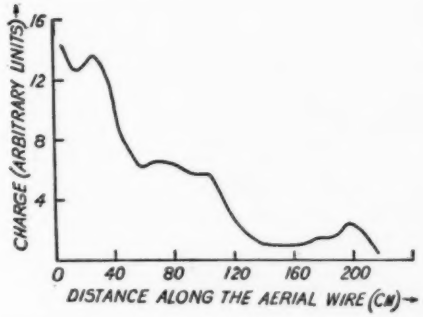


FIG. 7(c). Charge distribution along the length of the conical helix shown in Fig. 2. Base is close to the "ground" and feed is at the apex. Excitation frequency, 300 Mc/sec.

Or, rearranging the terms,

$$E_\phi = E_0 \sum_{m=1}^{N} a_m I_m \exp(j\psi_{\theta m}) \sum_{n=-\infty}^{n=+\infty} j^n J_n(d_\theta) \times \int_{-\pi}^{+\pi} \Big[\{\cos\phi(1-k_m\phi) - k_m \sin\phi\} \times (1 - jH\phi^2) \exp\{j(\mathfrak{C}_{\theta m} - g_m + n)\phi\} \Big] d\phi. \quad (18)$$

Introducing the exponential forms of $\cos\phi$ and $\sin\phi$ and putting,

$$\left.\begin{aligned} x &= v_{\theta m} - g_m + n + 1 \\ y &= v_{\theta m} - g_m + n - 1 \end{aligned}\right\}, \quad (19)$$

and also neglecting second-order terms,

$$E_\phi = \frac{E_0}{2} \sum_{m=1}^{N} a_m I_m \exp(j\psi_{\theta m}) \sum_{n=-\infty}^{n=+\infty} j^n J_n(d_\theta) \times \int_{-\pi}^{+\pi} \Big[(1 + jk_m - k_m\phi - jH\phi^2) \exp(jx\phi) + (1 - jk_m - k_m\phi - jH\phi^2) \exp(jy\phi) \Big] d\phi \quad (20)$$

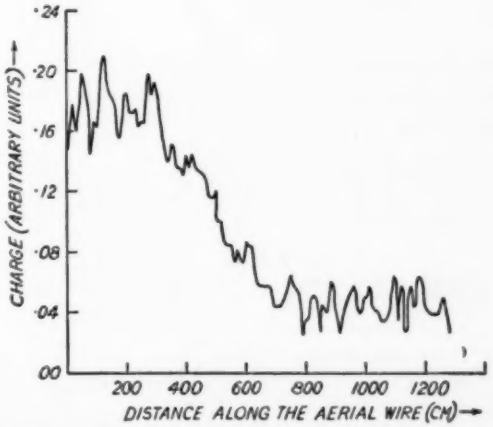


FIG. 8. Charge distribution along the length of the conical helix shown in Fig. 2. Base is close to the "ground" and feed is at the base. Excitation frequency, 200 Mc/sec.

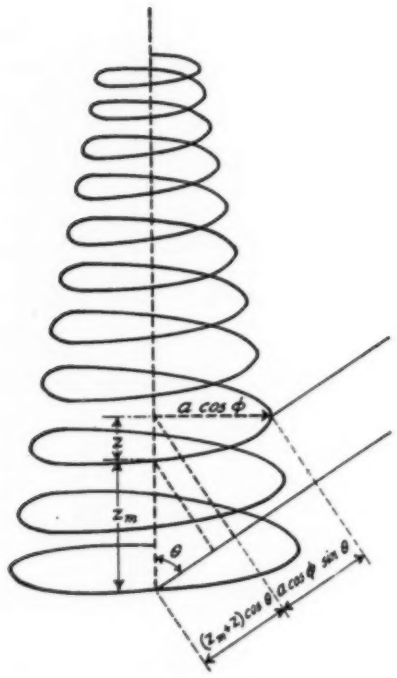


FIG. 9. Illustrating the calculation of the distance, from a point in the far zone, of an element ds of a loop of the helix with reference to the distance of the same point from the origin.

$$= E_0 \sum_{m=1}^{N} a_m I_m \exp(j\psi_{\theta m}) \sum_{n=-\infty}^{n=+\infty} j^n J_n(d_\theta)$$

$$\times\Bigg[\Bigg\{(1+jk_m)\frac{\sin x\pi}{x} - k_m\left(-j\pi\frac{\cos x\pi}{x} + j\frac{\sin x\pi}{x^2}\right) - jH\left(\pi^2\frac{\sin x\pi}{x} + 2\pi\frac{\cos x\pi}{x^2} - \frac{2\sin x\pi}{x^3}\right)\Bigg\}$$

$$+\Bigg\{(1-jk_m)\frac{\sin y\pi}{y} - k_m\left(-j\pi\frac{\cos y\pi}{y} + j\frac{\sin y\pi}{y^2}\right) - jH\left(\pi^2\frac{\sin y\pi}{y} + 2\pi\frac{\cos y\pi}{y^2} - \frac{2\sin y\pi}{y^3}\right)\Bigg\}\Bigg]. \quad (21)$$

We next obtain the θ-component of the electric field at the distant point. Substituting Eq. (7) in Eq. (5), E_θ is given by

$$E_\theta = -\frac{j\omega\mu_0}{4\pi R_0}\sum_{m=1}^{N}\int_{-\pi}^{+\pi}\Bigg[\cos\alpha\Bigg\{\cos\theta\left(\sin\phi - \frac{k_m\sin\phi}{1-k_m\phi}\right) + \tan\alpha\cdot\sin\theta\Bigg\}\cdot\exp\left(-j\frac{\omega}{c}R_m\right)\cdot I_m\exp(-jv_m)$$

$$\times\exp\left\{j\frac{\omega a_m\sec\alpha}{v_m}\left(\phi - \frac{k_m\phi^2}{2}\right)(1+h_m\phi^2)\right\}$$

$$\times a_m\sec\alpha(1-k_m\phi)\Bigg]d\phi. \quad (22)$$

Proceeding as before this expression can be simplified and E_θ is given by,

$$E_\theta = E_0\sum_{m=1}^{N} a_m I_m\exp(j\psi_{\theta m})\sum_{n=-\infty}^{n=+\infty} j^n J_n(d_\theta)$$

$$\times\Bigg[j\cos\theta\Bigg\{(jk_m-1)\frac{\sin x\pi}{x} + k_m\left(-j\pi\frac{\cos x\pi}{x} + j\frac{\sin x\pi}{x^3}\right) + jH\left(\pi^2\frac{\sin x\pi}{x} + 2\pi\frac{\cos x\pi}{x^2} - \frac{2\sin x\pi}{x^3}\right) + (1+jk_m)\frac{\sin y\pi}{y} - k_m\left(-j\pi\frac{\cos y\pi}{y} + j\frac{\sin y\pi}{y^2}\right) - jH\left(\pi^2\frac{\sin y\pi}{y} + 2\pi\frac{\cos y\pi}{y^2} - \frac{2\sin y\pi}{y^3}\right)\Bigg\}$$

$$+2\tan\alpha\sin\theta\Bigg\{\frac{\sin u\phi}{u} - k_m\left(-j\pi\frac{\cos u\pi}{u} + j\frac{\sin u\pi}{u^2}\right) - jH\left(\pi^2\frac{\sin u\pi}{u} + 2\pi\frac{\cos u\pi}{u^2} - \frac{2\sin u\pi}{u^3}\right)\Bigg\}\Bigg], \quad (23)$$

where $u = e_{\theta m} - g_m + n$.

The expressions for E_ϕ and E_θ are long, but some of the terms are quite small compared with the others. In the series

$$\sum_{n=-\infty}^{n=+\infty} j^n J_n(d_\theta),$$

the zeroth term is the largest and all others, except $n = \pm 1$, can be neglected as they are less than 1 percent. Of the terms in the parenthesis, $\sin x\pi/x$, $\sin y\pi/y$, $\sin u\pi/u$ are by far the largest terms and of these $\sin x\pi/x$ term is very large.

In Fig. 10 the experimental and theoretical polar diagrams for a conical helix (base nearer the ground), fed at base, are compared for 200 Mc/sec excitation.

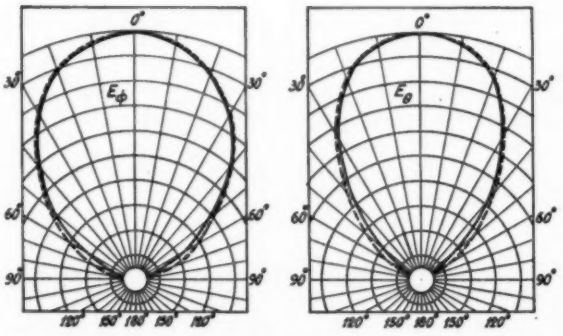


FIG. 10. Calculated (dotted lines) and observed (solid lines). E_ϕ and E_θ radiation patterns for the case shown in Fig. 3(c).

Only the first-order terms have been retained in the calculation. The phase velocity has been obtained from the following relation as used by Kraus:[7]

$$L/p=\lambda+(\lambda/2N)+Z, \tag{24}$$

where $p=v/c$, $L=$length of one turn, $Z=$the distance between the turns for the conical helix. The actual phase velocity for the conical helix may be slightly different from that given by Eq. (24) because all the turns are not of same diameter.

IV. CONCLUDING REMARKS

We note that while in a circular helix the axial mode of radiation can be maintained over a frequency range of about one octave, that in a conical helix can be maintained over a much wider frequency range by suitably varying the diameter. For a wide band conical helix the axial mode of radiation is, however, likely to be mixed with the normal mode in the low side of the frequency band. The relative magnitudes of the axial and the normal radiations depend on the pitch angle, the degree of tapering, and also on the feed position.

The conical helix can be modified in several ways for special applications. Some of these variations are shown in Fig. 11.

In Fig. 11(a) a conical helix in which the radius is constant in the central part is shown. This will have a sharper radiation pattern in the middle of the frequency band over which it can be used. The conical helix as shown in Fig. 11(b) is suited for balanced feeder. If the pitch angle is zero the conical helix reduces to a spiral as shown in Fig. 11(d). With appropriate delays at suitable intervals along the length as shown in Fig. 11(e), the spiral will radiate most of the energy in a direction perpendicular to the plane for a wide frequency range. Figure 11(e) shows a combination of spiral and conical helix which will have large band width with small height of the aerial above the ground.

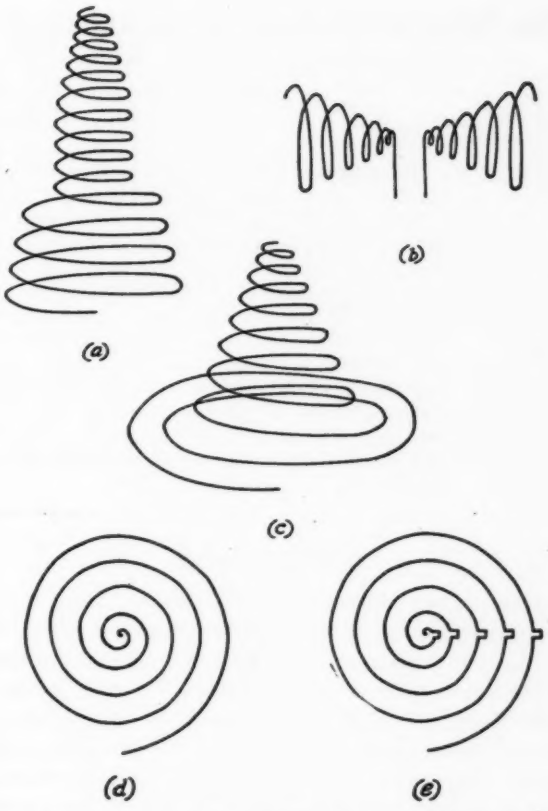


FIG. 11. Some modifications of the conical helix.

This type of aerial may be useful for ionospheric investigations.

ACKNOWLEDGMENT

The above investigations were carried out in the Institute of Radio Physics and Electronics, Calcutta University. The author wishes to express his sincere gratitude to Professor S. K. Mitra for constant encouragement and helpful advice during the progress of the work.

[7] J. D. Kraus, Proc. Inst. Radio Engrs. 37, 263 (1949).

The Effect of Pressure on the Tensile Properties of Several Metals and Other Materials

P. W. BRIDGMAN
Lyman Laboratory of Physics, Harvard University, Cambridge, Massachusetts
(Received October 31, 1952)

Tension tests to fracture under superposed hydrostatic pressure reaching to nearly 30 000 kg/cm² have been conducted on five polycrystalline metals, Ni, Ta, Cb, Mo, and W, two brittle single metallic crystals, Sb and Cu_5Zn_8, and two brittle amorphous materials, "Melmac 404" and B_2O_3 glass. All these materials exhibit ductility at the highest pressure; the largest increases are for the polycrystalline metals of which the largest is W, which breaks brittlely at atmospheric pressure but with 81 percent reduction of area at the highest pressure. The shape of the strain hardening curves varies among the five metals from abrupt rise with long tailing off to long gradual rise with comparatively abrupt falling off, in the order W, Mo, Ta, Cb, and Ni. The character of the fracture varies continuously with pressure and is different for the different metals. Of the brittle single crystals Sb has a reduction of area of 55 percent at the highest pressure. The plastic, Melmac 404, gives the smallest increase of ductility, a reduction of area of only 14 percent at the maximum. This small value is doubtless associated with the absence of internal crystal slip planes. B_2O_3 glass gave a reduction of 87 percent; this is largely a phenomenon of plasticity, but there is an appreciable component of viscosity. Both amorphous materials showed an increase of density after pulling under pressure instead of a decrease.

INTRODUCTION

IN this paper previous investigations[1] of the effect of hydrostatic pressures up to 30 000 kg/cm² are extended to a number of new materials. The majority of the previous work was done on various grades of steel. These all exhibited qualitatively similar effects. One of these effects was a very great increase, under pressure, in ductility, that is, the strain at fracture defined as $\log_e A_0/A_f$. Here A_0 is the initial area of the cross section and A_f the final area at fracture. Under the higher pressures, reductions of area so great as to be no longer measurable are obtained, the tensile specimen being pulled out practically to a point before fracture. This increase of ductility is known to be exhibited by other materials also, and a number of examples have been already described. A question examined in this paper is whether types of material different from those yet examined will also show the phenomenon. Several such materials are examined here. Steel exhibits, in addition to its great increase of ductility, at least two other characteristic effects, namely, a strain hardening which increases linearly with pressure, indicating by extrapolation the possibility of indefinitely high tensile strengths under sufficiently high pressures, and a linear dependence of the strain at fracture on the supporting pressure. These two linear relations are independent of each other. A question to be examined was whether other materials will also show these two linear relations. There had been some indication that, although the increase of ductility might be a rather common effect exhibited by many materials, the other linear relations were rather special to steel. Another question for examination was whether the empirical connection found between the reduction of area and the profile of the contour of a tensile specimen for steel[2] applies to other metals also. Choice of metals for the present investigation was somewhat limited by the instrumental requirement that they be in the same strength range as steel, a requirement imposed by the somewhat low sensitivity of the device with which tension is measured under pressure. This rules out the softer metals, which would doubtless be of equal interest with those measured here. It is also required that the metal be obtainable in the form of rod of sufficient length to yield a number of specimens, since the specimens must be similar. The five metals examined in detail here, namely, nickel, tantalum, columbium, molybdenum, and tungsten, were the only metals commercially obtainable which satisfied these requirements.

In addition to the conventional tests at atmospheric pressure, two types of measurement were made. The first was a determination of the effect of pressure on ductility, and it involved pulling the specimen to fracture under several pressures spaced over the range and plotting the reduction of area at fracture as a function of pressure. These ductility measurements often did not extend over the entire pressure range, because several of the metals neck down to an immeasurably small section under pressures much below 30 000 kg/cm². The second sort of measurement was of the strain hardening curve. This involved pulling a specimen to a predetermined reduction of area at an assigned pressure, removal from the pressure apparatus, measurement of the reduction of area, followed by replacement in the pressure apparatus and further pulling at the same pressure. This was repeated several times and yielded a curve of reduction of area *versus* the corresponding tensions for the particular fixed pressure. The entire procedure was repeated with other specimens at other fixed pressures, yielding finally a set of strain hardening curves as a function of pressure. In order to obtain the true flow stress it was necessary to correct the measured tensile load for nonuniformity of stress across the section, and this demanded a deter-

[1] P. W. Bridgman, *Studies in Large Plastic Flow and Fracture* (McGraw-Hill Book Company, Inc., New York, 1952).
[2] Reference 1, p. 26.

mination of a/R, where a is the radius of the cross-sectional area, and R is the radius of curvature of the contour. The necessary measurements were made after each release of pressure.

The same apparatus and experimental procedure was used as in previous work on steel; these have been fully described.

Isolated measurements were made on single-crystal gamma-brass and on single-crystal antimony as perhaps extreme examples of brittle materials that might possibly acquire measurable ductility under pressure. One plastic, "Melmac 404," was tried with the idea that there might be classes of materials which instead of becoming ductile under pressure become more brittle and glass-like. Finally boric anhydride glass was tried with the idea that if true viscosity continues to be a property of glass down to room temperature, measureable deformations ought to be producible in cold glass. This glass, because of its unusually low normal softening point, seemed the most promising material to exhibit such an effect.

THE DETAILED RESULTS

Nickel

This material was obtained through the courtesy of Mr. F. B. Foley and the International Nickel Company. It was furnished in the form of a bar 0.5 inch in diameter, annealed, and stated to be of 99.4 percent purity. Since the specimens required were only 0.25 inch in diameter, in order not to waste material the bar was flattened by a light forging at a red heat and then annealed by holding at 1600°F for 30 minutes with slow cooling in the furnace; the annealing process is described in Technical Bulletin T-20 of the International Nickel Company.

Measurements were made on nine specimens, three of these at atmospheric pressure. Nickel is normally very ductile, the strain at fracture at atmospheric pressure being 1.76, corresponding to a reduction of area of 83 percent. Fracture under a pressure of 7700 kg/cm² occurred at a strain of 2.58 (reduction of area of 92.5 percent). There was much uncertainty in this determination because of irregularities in the section at fracture. At higher pressures the specimens pulled down to irregularly arranged knife edges before fracture so that it became meaningless to attempt to assign any area at fracture and the ductility cannot be given at pressures higher than 7700.

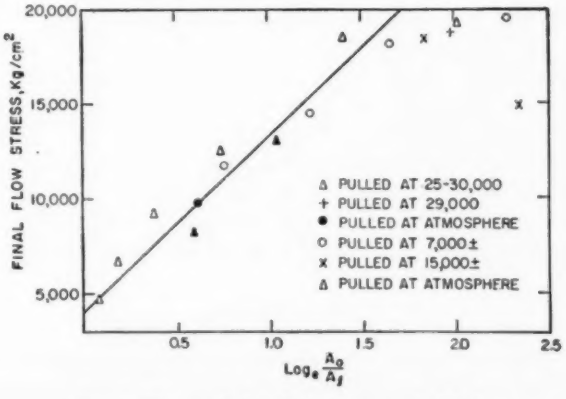


FIG 1. The strain hardening curve of nickel.

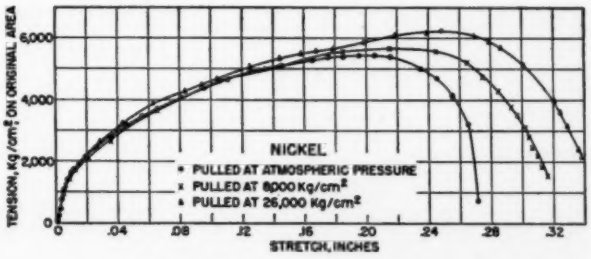


FIG. 2. Curves of load *versus* stretch for nickel at several pressures.

The strain hardening data are shown in Fig. 1. In this diagram are data for six different specimens, pulled at different pressures, as indicated. To a first approximation the data for the different specimens fall on the same curve, irrespective of the pressure of pulling. To a second approximation, however, the points for the specimens pulled at the higher pressures lie higher. The same effect was found with steel. Over most of the range, up to a strain of 2, the strain hardening curve is approximately linear. Above this, it shows a tendency to turn down, which means that at extreme strains there is a softening rather than a hardening under an increment of strain. This is probably the result of actual material damage to the specimen at the very high strains, perhaps because of lack of homogeneity in the original material. This is consistent with the appearance of the fracture.

Strain hardening in nickel is a slow process and the maximum in the curve of stretch *versus* total load which marks the beginning of necking is unusually long deferred. In Fig. 2 are shown three curves of stretch *versus* total load (or what is the same thing, tension in kg/cm² on the original area) for three different pressures. These curves rise higher and the maximum is displaced to greater elongations for the higher pressures. The maximum total load increases from 5400 kg/cm² at atmospheric pressure to 6200 at a pressure of 26 000 kg/cm², an increase of 15 percent, or 5.7 percent for 10 000 kg/cm². This is of the same order of magnitude as found for steel.

The original elastic limit of this nickel was not marked well enough to justify the attempt to find the effect of pressure on it.

In order to find the correction factor for reducing "true" tensile stress (that is, total load divided by actual neck area) to "flow" stress, it was necessary to determine a/R. The correction factor for converting true stress to flow stress as given by the measured a/R was consistently higher than the "empirical" factor

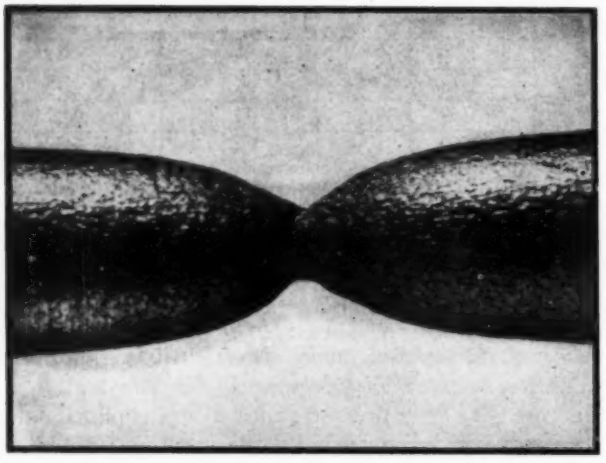


FIG. 3. Photograph of a tension specimen of nickel pulled under a mean hydrostatic pressure of 16 000 kg/cm².

determined by using the empirical relation found for steel between strain and a/R. The excess was of the general order of 7 percent. In other words, the empirical relation found for steel between strain and a/R does not hold for nickel. The direction of difference means a smaller curvature in the contour for nickel than for steel; this is consistent with the delayed appearance of necking.

The fractures were never as regular as for steel. At low pressures the general character was a rough cup-cone, with more irregular edge and bottom than for steel. At high pressures the tendency was to draw out to a chisel edge. The flow pattern on the surface remote from the fracture was a general roughening of the surface, without longitudinal streaks. The surface texture becomes finer for the higher pressure. In Fig. 3 is shown an unfractured specimen pulled at 16 000 kg/cm².

Tantalum

The specimens were made from annealed ¼-inch rod of 99.9 percent purity obtained from the Fansteel Company. Seven specimens were used; three were pulled to fracture, and the strain hardening curves were determined with the four others.

Tantalum is already so ductile at atmospheric pressure that it was not possible to measure the increase of ductility under pressure. At atmospheric pressure one specimen pulled down to a knife edge at fracture, and at higher pressures the two opposing fragments were always drawn out to sharp edges, of contorted shapes.

The strain hardening curves for several different pressures are shown in Fig. 4. Within experimental error the flow stress increases linearly with the strain up to strains of 2. Measurements were not pushed to higher strains because of increasing geometrical irregularities at the neck. Figure 4 shows a very definite effect of pressure on the strain hardening, the flow stress being greater for the same strain at higher pressures, as would be expected. The rate of increase of strain hardening is only one-half as great for tantalum as for nickel.

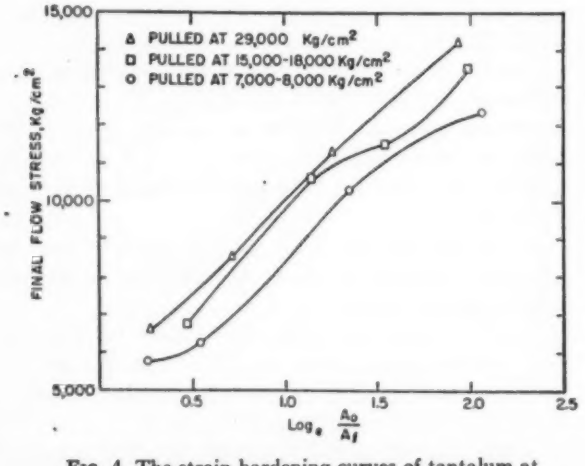


FIG. 4. The strain hardening curves of tantalum at several pressures.

The conventional tension curve, stress on the original area *versus* total stretch, is shown for several pressures in Fig. 5. The general nature of the curves is strikingly different from those for nickel. Here the curve rises to an early maximum, which is followed by a long gradual tapering off. This means early onset of necking. Consistent with this is the fact that the "correction factor" determined by using the empirical relation between strain and a/R that was valid for steel is now too high, instead of too low, as for nickel. That is, the shapes of the neck of tension specimens of nickel and tantalum diverge in opposite directions from that of steel.

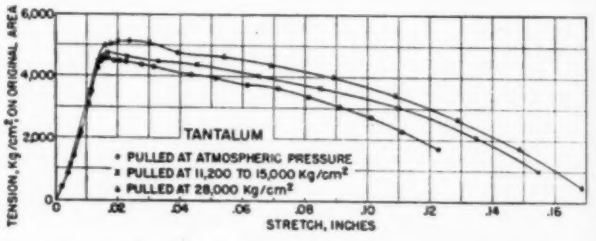


FIG. 5. Curves of load *versus* stretch for tantalum at several pressures.

At atmospheric pressure the fracture is a crudely developed crater with coarse granular structure. At higher pressures the structure becomes finer and the crater evolves into a knife edge arrangement, until at 28 000 kg/cm² the fracture is a star-shaped arrangement of three knife edges, as shown in Fig. 6. The surface structure remote from the fracture is strongly ribbed longitudinally, becoming finer at the higher pressures.

Columbium

The specimens were made from annealed ¼-inch rod of presumptive 99.9 percent purity obtained from the Fansteel Company. Columbium is if anything more

ductile than tantalum, so that only the strain hardening curves could be determined. For this, four specimens were used. Of these one was pulled to fracture in two stages at atmospheric pressure, one was pulled to fracture in one stage at a pressure starting at 16 500 and ending at 28 800 kg/cm², one was pulled in six stages at a mean pressure in the neighborhood of 28 000, and one was pulled in one stage at 29 000 to a reduction of area of 62 percent, from which pulling was continued to fracture at atmospheric pressure.

Curves of total stretch *versus* total load are shown in Fig. 7 for atmospheric pressure and 28 000 kg/cm². The maximum total load and the maximum elongation are greater under pressure, as is also the case for nickel and tantalum. The maximum is 4600 kg/cm² under 28 000 and 3900 at atmospheric, an increase of 6.4

FIG. 6. Photograph of a tensile specimen of tantalum pulled to fracture under a pressure of 28 000 kg/cm². The specimen pulled out to a point; the photograph is taken looking down on the point.

percent for 10 000 kg/cm². This is in the range for other metals.

The strain hardening curve was determined only for the highest pressure; it is shown in Fig. 8. The rate of increase of hardening is high compared with other metals. There is considerable uncertainty in this, however, because of rather unusually large irregularity in the cross-sectional area and in the contour, making uncertain the correction factor for converting from average actual stress to flow stress. The hardening curve is shown only as far as a strain of 1.54. Pulling was actually continued to a strain of 2.47, but the correction factor could not be determined at this strain. It would appear probable, however, that strain hardening does not continue to as high a strain as this, but deterioration of the material sets in, because if average load over the

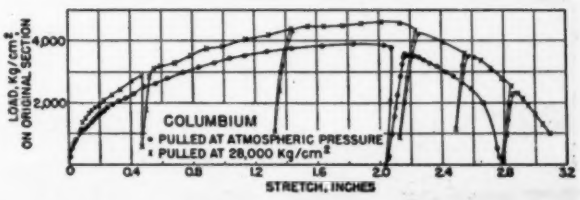


FIG. 7. Curves of load *versus* stretch for columbium at two pressures.

section is plotted against strain instead of flow stress, a curve will be obtained which turns over at the highest strains. The average stress on the section was 11 700 kg/cm² at a strain of 1.54 and 12 000 kg/cm² at 2.47.

The fracture at atmospheric pressure is cup-cone with fine granular structure. The strain at atmospheric at fracture is 2.23 (91 percent reduction of area); it is unusual to find a cup-cone fracture with such great reduction. At higher pressures the character of the fracture changes, until at 28 000 the fracture is a star of three chisel edges, very much like that of tantalum. The surface flow structure is a combination of pitting and longitudinal ribbing. Figure 9 shows a specimen pulled at 28 600.

The maximum load, that is, the beginning of necking, occurs at a strain intermediate between that of nickel and tantalum, more nearly as for steel. It is a consequence that the correction factor for converting average stress to flow stress agrees fairly well with the factor determined for steel. The difference is in the direction corresponding to an earlier beginning of necking in columbium than in steel.

The specimen pulled under 29 000 to a strain of 0.96 and then repulled at atmospheric pressure did not retain the full amount of hardening which it had acquired under pressure. Under 29 000 it was sustaining a load of 3450 kg/cm², calculated on the original section, and when repulled at atmospheric pressure the

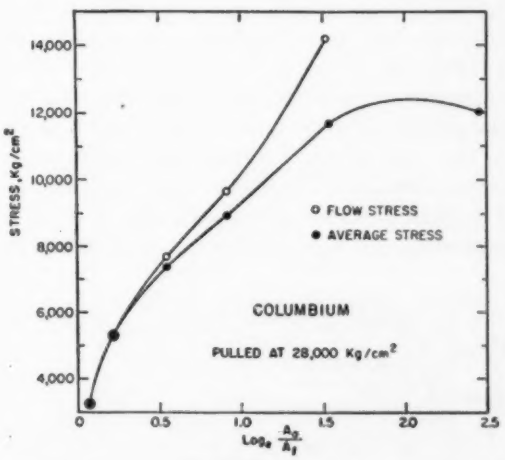


FIG. 8. Flow stress and average stress of columbium pulled at 28 000 kg/cm² as a function of natural strain. The flow stress curve is the same as the "strain hardening" curve.

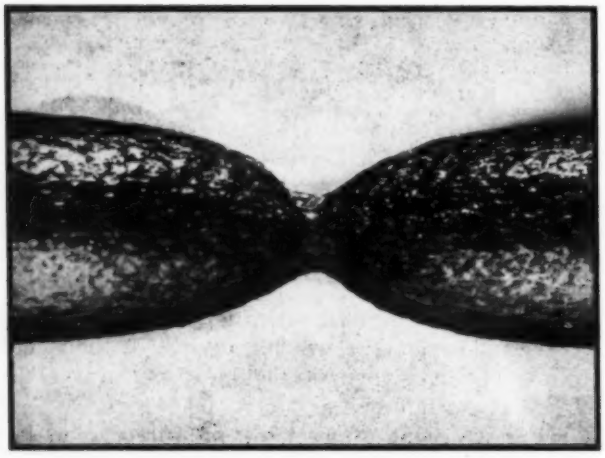


FIG. 9. Photograph of tension specimen of columbium after the sixth stretching at 28 600 kg/cm².

highest load it sustained was 3050 kg/cm². The pulling under pressure was, however, more effective in imparting hardening than a pulling at atmospheric pressure, for the load which the specimen sustained at atmospheric under an elongation equal to that under 29 000 was 2500 against 3050 above. The fracture at atmospheric of the prepulled specimen was cup-cone, very much like the fracture without prepulling.

Molybdenum

Five specimens were made from annealed $\frac{1}{4}$-inch rod obtained from the Fansteel Company. The purity was stated to be 99.9 percent. One specimen was pulled to fracture at atmospheric pressure in one stage, one was pulled to fracture in five stages at a mean pressure of 7500 kg/cm², one to fracture in five stages at 17 000, one to fracture in two stages at 28 000, and one not to fracture in three stages at 26 000 to 29 000.

The curves for stretch *versus* total load for three of these specimens are shown in Fig. 10. The general character of the curves is similar to those for tantalum —an abrupt rise to an early maximum, followed by a long tailing off. There is a notable increase of ductility with pressure; the atmospheric specimen fractured at an elongation of 0.065 inch at a total load 8 percent less than the maximum, whereas the elongation of the specimen fractured at 28 000 was 0.22 inch and the load to small to measure, certainly less than one percent of maximum. It was, however, not possible to draw

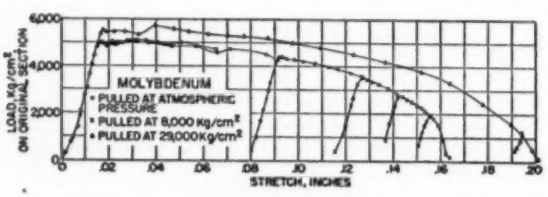


FIG. 10. Curves of load *versus* stretch for molybdenum at three pressures.

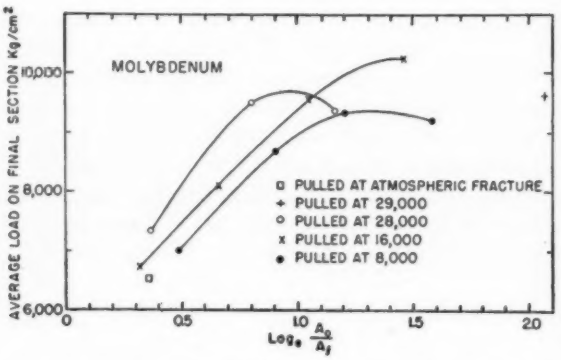


FIG. 11. Average load on final section ("true" stress) of molybdenum as a function of natural strain at several pressures.

the conventional ductility curve because of indeterminateness of the area at fracture, the distortion of the specimen being usually great.

In Fig. 11 are shown the average loads on the final section for the intermediate stages of pulling before the final fracture for the specimens pulled under pressure. The points corresponding to fracture of these specimens cannot be shown, the area at fracture being too indeterminate. The area at fracture of the atmospheric specimen was sufficiently defined to measure, however, and the corresponding point is shown. The figure shows in the first place a rather greater dependence than usual of the curves on the pressure of pulling, and in the second place a turning over of all the curves at the higher strains after an initial linear section. This turning over is without doubt due to deterioration of the material because of its initial unsatisfactory homogeneity, as was strongly suggested by the appearance of the specimens.

In Fig. 12 are shown the flow stresses as a function of strain for the specimens of Fig. 11 in the initial linear part of the curves before deterioration became important. The rate of strain hardening is usually high, and again there is an unusually large dependence on the pressure of pulling. How much of this is due to imperfect homogeneity of the specimen and how much is intrinsic to molybdenum would demand further investigation with other specimens.

The correction for reducing mean stress to flow stress in terms of the measured contour was in all cases smaller than the correction given by the empirical connection between contour and strain valid for steel. This discrepancy is in the opposite direction from that which would be expected from the early occurrence of necking shown in Fig. 10.

The fracture at atmospheric pressure was on a rough approximately plane surface perpendicular to the axis, without cup-cone effect. At higher pressures the fracture metamorphosed into a frazzled jagged pattern, reminicent of a piece of wood, tending at the highest pressures to fracture as chisel edges. The surface pattern at pressures high enough to support considerable flow

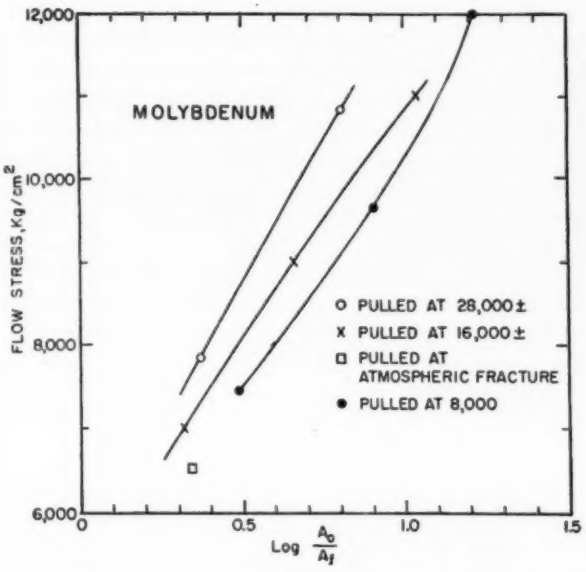


FIG. 12. Strain hardening curves of molybdenum at several pressures.

was strongly ribbed longitudinally, the ribs being frequently interrupted by shallow transverse fissures. Figure 13 shows the specimen after the fourth stretching at 17 000 kg/cm².

An experiment was made on one single crystal of pure molybdenum which was obtained through the courtesy of Professor Robert Maddin of The Johns Hopkins University. The original specimen was approximately 0.1 inch in diameter and ¾ inch long. Since it was not possible to fashion this into a conventional tension specimen, it was pulled by the copper sheath technique[3] with steel pull pieces at the ends. This technique is applicable only under comparatively high hydrostatic pressure.

The single crystal was pulled in two stages to fracture; the pressure at which fracture occurred was 22 000 kg/cm². It pulled down to a sharp chisel edge of width nearly the diameter of the original crystal, exactly like the familiar and classical fracture of single zinc crystals. The surface texture was only slightly roughened. The maximum load was reached at a stretch roughly one-third that at fracture, which means that necking begins later than with drawn rods. The maximum stress supported by this crystal was of the order of only 2000 kg/cm².

Tungsten

The specimens were cut from annealed ¼-inch rod of 99.9 percent purity provided by the Fansteel Company. At atmospheric pressure this tungsten rod is brittle and breaks with no measurable reduction of area. This means that the sharp corner at the head of the specimen has to be relieved by an unusually large fillet. Two specimens were lost by fracture at the corner of the

[3] Reference 1, p. 109.

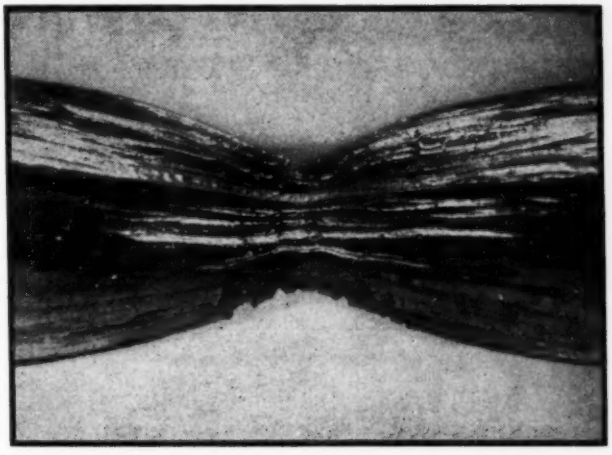


FIG. 13. Photograph of tension specimen of molybdenum after fourth pulling at a pressure of 17 000 kg/cm².

head before satisfactory dimensions for the fillet were found. Successful measurements were made on six specimens. Four of these were used in determining the effect of pressure on ductility, being pulled to fracture in a single stage at pressures of atmospheric, 7000, 16 000, and 27 800 kg/cm². The other two specimens were used to obtain the strain hardening curve and were pulled in four stages without fracture at mean pressures of 17 000 and 28 000.

In Fig. 14 is shown the total load as a function of the measured stretch for the four specimens pulled to fracture at successively higher pressures. The absolute strength of tungsten is of the order of twice that of the four preceeding metals. The maximum load increases with pressure, from 11 800 under 7000 kg/cm² pressure to 13 000 under 28 000, an increase of 4.5 percent for

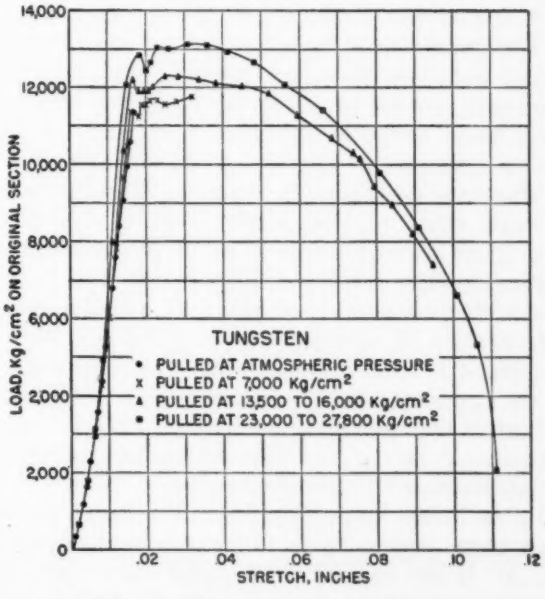


FIG. 14. Curves of load *versus* stretch for tungsten at four pressures.

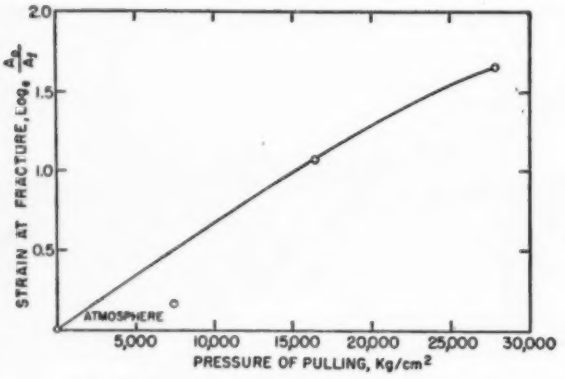


FIG. 15. Ductility of tungsten as a function of pressure of pulling.

10 000 kg/cm². This is of the same order of magnitude as for steel and the other metals, but rather on the small side.

The curves of Fig. 14 show distinctly the phenomenon of the second yield point, an effect which becomes more accentuated at the higher pressures. This is by far the best developed example of this phenomenon which has been found under pressure. Several grades of steel showed it, but with diminishing prominence under pressure. The effect for steel was so ill defined that no mention was made of it in reference 1. In the present investigation molybdenum exhibits the effect to a much less marked degree.

Figure 15 shows the effect of pressure on ductility; the pressure of pulling is indicated at the several points. The effect seems to be approximately linear in pressure, the low-lying point at 7000 kg/cm² probably being due to an inperfection in the specimen. The rate of increase of ductility with pressure is less than for the softer steels and is roughly the same as for a high carbon steel heat treated to a Rockwell C hardness of 56.

The strain hardening data for the two specimens pulled in stages, with two isolated points from other specimens, are shown in Fig. 16. It was not possible to carry these curves to as high strains as the ductility curves because irregularity of the contour at high strains made uncertain the correction for reducing from average stress to flow stress. Within experimental error there seems to be no effect of pressure on strain hardening. The relation with pressure appears to be linear except for the last point at which a marked deterioration of the specimen is indicated. This deterioration at high strains is in line with the appearance of the specimens, which showed various irregularities indicative of inhomogenities in the original material. Tungsten is by far the least satisfactory of the five metals measured here in this respect.

Up to 16 000 kg/cm² the fracture is on an approximate plane, perpendicular to the axis and finely granular. At 27 800 kg/cm² the fracture has assumed the irregular appearance shown in Fig. 17. An interesting example of partial fracture was obtained at 28 800, there being partial separation into jagged interpenetrating prongs with a superposed separation on longitudinal planes, one of which permitted clear passage of light through the specimen. This is in line with the deterioration at the upper end of the strain hardening curve already commented on. The surface slip texture was much finer than for the other four metals, and was a combination of a general roughing and a fine longitudinal ribbing.

The correction for the shape of the contour by which average stress is reduced to flow stress was, in general, less than that given by the empirical relation for steel. This means that a/R at the same strain is less for tungsten than for steel.

Gamma-Brass (Cu_5Zn_8)

The material was in the form of single crystals which were obtained through the courtesy of Dr. Cyril Smith of the Institute for the Study of Metals of the University of Chicago. It was furnished in the form of rods approximately 3 mm in diameter. It was necessary to handle it by the copper sheath technique already applied to single crystal molybdenum. This material is excessively brittle under ordinary conditions, and the chief interest was to find whether it would exhibit measurable ductility. Only one specimen was subjected to tension, a specimen which had previously been used[4] for the measurement of linear compressibility to 30 000 kg/cm². This was pulled in two stages at a mean pressure of 27 000 kg/cm². The experiment was terminated by accidental separation of the copper sheath, without apparent fracture of the specimen. However, after removal of the copper sheath at atmospheric pressure by filing, the specimen was found broken into four pieces, doubtless by an incidental result of the filing. The specimen showed a 10 percent reduction of area. Since this answered the fundamental question in the affirmative, namely, that measureable ductility is

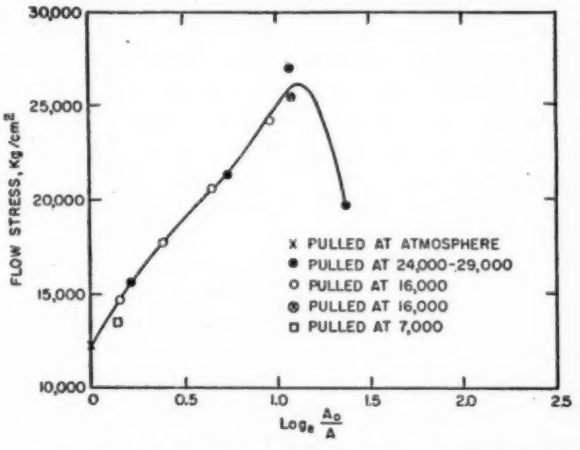


FIG. 16. Strain hardening of tungsten for specimens pulled at several pressures.

[4] P. W. Bridgmann, Proc. Am. Acad. Arts Sci. 77, 217 (1949).

imparted by pressure, the measurements were not pushed further. The relation between measured stretch and load on the original section is shown in Fig. 18. This is remarkable for the absolute load reached, 14 000 kg/cm², greater than for tungsten, and also remarkable for the approximately linear rise of load with stretch over an unusually wide interval. No appreciable necking was shown, as would be expected from Fig. 18, since the load did not pass through a maximum.

The fracture was very coarsely granular, irregularly perpendicular to the axis, with no suggestion of cup-cone.

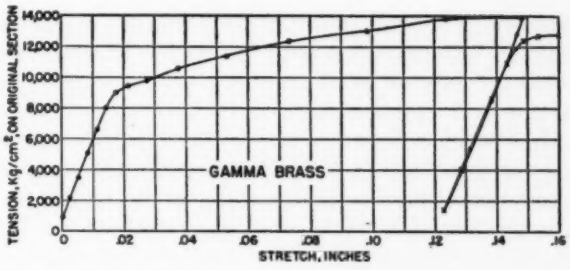


FIG. 18. Curves of load *versus* stretch for gamma-brass pulled under a pressure of 27 000 kg/cm².

Antimony

This was as single crystal, prepared from Kahlbaum "*K*" material left from my previous measurements[5] of

FIG. 17. Photograph of tension specimen of tungsten fractured by pulling under a pressure of 27 800 kg/cm².

resistance under pressure at low temperatures. It was approximately 3 mm in diameter and was mounted by the copper sheath technique.

It was pulled to fracture at a pressure varying from 22 400 to 26 000 kg/cm². The relation between total load and extension is shown in Fig. 19. The maximum load was 1520 kg/cm², and the load at fracture (on the original area) was 1160 kg/cm². The ratio of initial area to area at fracture was 2.22, equivalent to a natural strain of 0.800. The average final load ("true stress") at fracture was 2560 kg/cm².

[5] P. W. Bridgman, Proc. Am. Acad. Arts Sci. 68, 105 (1933).

The fracture was on a mirror-like plane, perpendicular to the axis, and appreciably displaced from the narrowest part of the neck.

Melmac 404

This plastic is melamine formaldehyde, with no filler, an experimental product which was obtained by the author several years ago from the American Cyanamid Company for measurements of compression to 40 000 kg/cm². Of all 40 plastics measured at that time,[6] Melmac 404 has the least compression under 40 000, the volume decrease being approximately 17 percent. Under ordinary conditions it is brittle and breaks easily. For this reason it seemed best suited to my purpose, which was to find, if possible, a plastic which retained a large degree of its initial brittleness under pressure. Perhaps one's first inclination would be to think that under pressure a "plastic" would become indefinitely ductile, considering the great increase of ductility exhibited by the metals. But on second thought it would seem that the mechanism of slip on crystal planes is not present in a plastic, which must be more glass-like in structure, and which therefore may perhaps retain considerable brittleness under pressure.

Three specimens of Melmac 404 were pulled, all by the copper sheath technique. The first was pulled in a

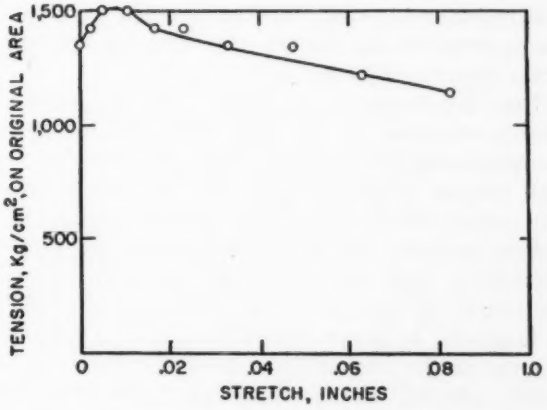


FIG. 19. Load *versus* stretch curve of single crystal antimony pulled under a maximum pressure of 26 000 kg/cm².

[6] P. W. Bridgman, Proc. Am. Acad. Arts Sci. 76, 71 (1948).

single stage, not to fracture, under a maximum pressure of 29 000 kg/cm²; the second was pulled to fracture at a final pressure of 24 600, and the third was pulled to fracture at a final pressure of 7000 kg/cm².

The specimen fractured under 24 600 broke *in shear* on a plane inclined at 45° to the longitudinal axis with a reduction of area of 17.6 percent, corresponding to a natural strain of 0.193. It would appear, therefore, that this plastic does retain an unusually great degree of brittleness under pressure. Initial plastic yield began at a load of approximately 4000 kg/cm² calculated on the original area. After the initial yield load and length increased linearly together to fracture which occurred at a load of 6060 kg/cm² on the original area, or 7330 kg/cm² on the actual area. There was no necking, as of course there should not have been because load did not reach a maximum, but the decrease of area was distributed uniformly along the length. This pulling was accompanied by a paradoxial decrease of volume (increase of density) to 0.944 initial. The purely elastic effect of tension is of course an increase of volume. Doubtlessly the decrease of volume produced by the hydrostatic pressure, which is numerically greater than

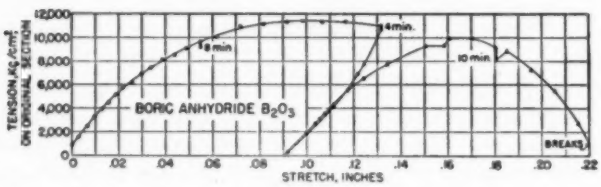


FIG. 20. Load *versus* stretch for boric anhydride glass pulled under pressures varying from 23 000 to 29 000 kg/cm².

the volume increase produced by the tension, got frozen in by some orientation effects of the pulling.

The specimen fractured under a pressure of 7000 kg/cm² and supported a load on the original section of only 3450 kg/cm². Fracture occurred very soon after the first perceptible plastic yield and with no measureable reduction of area. The fracture was a mirror-like plane, perpendicular to the axis, strongly reminiscent of the fracture of glass.

These measurements on Melmac 404 were unlike any hitherto reported in the large variations of Young's modulus in the elastic range. This was a variation both with respect to pressure and with respect to tension. The initial stiffness of the two specimens pulled at the highest pressures was three times as great as the initial stiffness of that pulled at 6000 to 7000. The final stiffness of the unruptured specimen pulled at 29 000, on beginning to release tensile load, was at least twice as great as the final stiffness when the load was completely relaxed. Taken altogether, here is a sixfold variation of Young's modulus with pressure and tension. No doubt a substantial part of this would have been found to be tied up with hysteresis effects if a complete tension cycle could have been taken. The present method is not well adapted to give the absolute value of Young's modulus, but it appeared fairly certain that the initial modulus at high pressures is less by a factor of at least three than the modulus for gamma-brass.

Boric Anhydride Glass, B_2O_3

This is material left from the work of Birch and Dane[7] on the effect of pressure at elevated temperatures on viscosity. The compressibility of this identical material has been determined[8] to 100 000 kg/cm².

Successful measurements were made on one specimen, selected for freedom from visible bubbles or other flaws, by the copper sheath technique. It was subjected to two pullings, the second to fracture, at pressures varying from 23 000 to 29 000 kg/cm². The relation between load and elongation is shown in Fig. 20. Deviation from linearity between elongation and load is initiated gradually, with no well-defined elastic limit. On the first pulling, load was continued beyond the maximum, with presumptive corresponding beginning of necking. This was confirmed by the appearance after release of pressure. The permanent increase of length after the first stretching was 7.5 percent.

On first release of tension the relation between load and extension departs notably from linearity, the slope decreasing by a factor of approximately two during release, with the greater slope and therefore the greater effective Young's modulus at the maximum load. The slope at small loads on initial application of load corresponds to an effective Young's modulus of roughly 500 000 kg/cm². At small loads on first release the effective Young's modulus is roughly ten percent less than this; more than half the difference of slope exhibited in Fig. 20 is due to permanent change of geometrical dimensions.

On second application of load up to roughly half the previous value the relation between load and extension is approximately linear, but with perceptible hysteresis in the usual direction. Beyond half-load the unloading curve is markedly deviated from in the direction of smaller load for the same extension. This means a strain softening instead of a strain hardening as a result of the first stretching. This is without doubt to be ascribed to the true viscous flow which evidently occurs to a small degree as an accompaniment of the conventional "plastic" behavior. This viscous flow is manifested as a relaxation of load with time at constant extension. Rough measurements were made of relaxation in times less than ten minutes at extensions of 0.05, 0.13, and 0.18 inch. There was no variation in the relaxation rate greater than would be expected from changes of dimensions in this range. The order of the relaxation was 100 kg/cm² per minute.

The second pulling was continued to fracture, the latter end of the load-extension curve being similar to

[7] E. B. Dane, Jr., and F. Birch, J. Appl. Phys. **9**, 669 (1938).
[8] P. W. Bridgman, Proc. Am. Acad. Arts Sci. **76**, 67 (1948).

that for a metal tensile specimen except for the greater relaxation. The final reduction of area at the neck was by a factor of 7.72, corresponding to a natural strain at fracture of 2.04, or a reduction of area of 87 percent.

Immediately after the fracture of this specimen pressure was released rapidly, and the specimen removed and placed in dry ice to retard any internal changes. From fracture to dry ice was 20 minutes at room temperature. The copper sheath was then removed by filing, at dry ice temperature. In the next 45 minutes the fragments were transported to the Laboratory of Arthur D. Little, Inc., and the density was most kindly measured by Dr. I. Simon, who had been doing extensive work of this kind on other sorts of glass. The final density determination was at 31°C and occupied about 7 minutes. The method was a floatation method with subsequent determination of the density of the floatation liquid in a pyknometer. The mean density of one-half of the broken specimen was 1.977 and that of the other 1.980. The density of the control, virgin B_2O_3, was 1.832. The density of the fractured specimen was thus increased by an average of 7.96 percent. There are two effects operative here; pure hydrostatic pressure, which increases density, and tension, which decreases density. Hydrostatic pressure is not usually thought of as producing permanent changes of density; in fact, it does not produce such changes for perfect crystalline material, as I have often checked. Glass, however, seems to be a different class of material, and I have recently found[9] permanent changes of density produced by high pressure. The increase is by no means linear in pressure and does not appear at all until a critical threshold pressure is reached. For some of the harder glasses the critical pressure seems to be as much as 50 000 kg/cm², but for this borax anhydride glass it appears to be considerably lower. Thus I have found a permanent change of density of 3.5 percent after exposing to 40 000 kg/cm² (experiments unpublished) in my standard apparatus[6] for determining compressibility to 40 000, pressure being transmitted by indium. In the experiment above, the stretching was conducted at a lower pressure, 29 000 maximum, and the permanent increase of density was twice that found under 40 000. The explanation is probably to be found in the action of the tension. During the very considerable longitudinal flow under tension, the molecules of the glass are to be thought of as sliding over each other, affording opportunity for the pressure to close pack the molecules more effectively than when the deformation is approximately only a hydrostatic compression.

Tension tests under pressure were also made on a second specimen, with the particular object of studying more carefully the relaxation phenomena. This experiment was not successful, however, but was terminated by failure of the copper sheath after a permanent extension of about 10 percent, and before necking started, at a pressure of 28 000 kg/cm². The density of this specimen after pulling, determined with the same precautions as to temperature as before, was 1.905. A blank on a piece of glass subjected to 29 000 twice, without tension, gave a density of 1.939, an increase of 5.9 percent against the 3.5 found before for 40 000. Part of the difference may be due to a failure of pressure in the indium to be truely hydrostatic, or perhaps more probably to the duration of the pressure. The low figure 1.905 just mentioned may be explained partly as an intrinsic effect of tension and partly different duration. The matter evidently should be studied further, but of the existence of the effects there can be no doubt.

[9] P. W. Bridgman, Proc. Am. Acad. Arts Sci. 81, 170 (1952).

SUMMARY

In this paper previous investigations of the effect of pressure on tensile properties have been extended to nine new materials: five polycrystalline metals, two single crystals, and two amorphous substances, including one plastic and one low melting glass. In general, the previous result holds, namely, that pressure tends to increase ductility, sometimes by very large amounts. There seem to be two classes of material with regard to the order of magnitude of the increase, namely, crystalline materials, including polycrystalline materials, and noncrystalline materials. The effect is much smaller for the latter.

The polycrystalline metals studied here, nickel, tantalum, columbium, molybdenum, and tungsten, all have intrinsic strengths of the same order as steel and exhibit somewhat similar effects. There are very large increases of ductility. Strain hardening at first increases markedly with pressure. The strain hardening of several of the metals does not continue to increase indefinitely, however, as does that of steel, but the curve may turn over with eventual strain softening rather than hardening. The appearance of the specimens suggests that such softening is not an intrinsic property of the metal but is a result of insufficient homogeneity in the original material. Besides increasing strain hardening, pressure raises the whole general level of the stress-strain diagram by an amount roughly the same as that for steel, that is, by something of the order of 5 percent per 10 000 kg/cm². The strain hardening curves of the different metals differ greatly in general appearance: They may rise abruptly, with a sharp turn-over followed by a long tailing off, or they may rise gradually to a flat maximum followed by a rather abrupt drop. Tungsten exhibits the phenomenon of a second yield point under pressure, the effect becoming more marked at higher pressures. Molybdenum shows the effect to a less degree. All the metals exhibit progressive changes in the character of fracture: in general the original cup-cone disappears and the fracture occurs on jagged prongs or on shear planes. At the same time the general texture of the fracture becomes finer.

The two unicrystalline materials, originally excessively brittle, acquired ductility under pressure. The experiment on gamma-brass was terminated by an accident before the maximum amount of ductility that could be imparted was determined. Antimony permitted a 55 percent reduction of area.

One of the amorphous materials, the plastic Melmac 404, broke at the highest pressure with a reduction of area of 14 percent. This is much less than that of antimony and suggests that the occurrence of internal slip planes may be important if pressure is to impart a large increase of ductility. Boric anhydride glass permitted a reduction of area of 87 percent. This very large effect is doubtlessly connected with the low melting point of this glass and the retention at room temperature of some degree of true viscosity. Both the plastic and the glass showed increases of density after pulling under pressure of the order of several percent. It now appears probable that in general noncrystalline substances are capable of permanent increases of density after exposure to pressures above some critical value and that pressure is more effective in imparting such an increase of density if the substance is being plastically deformed, as by tension, at the same time that pressure is acting.

I am much indebted to Mr. L. H. Abbot for making most of the readings of this paper.

JOURNAL OF APPLIED PHYSICS VOLUME 24, NUMBER 5 MAY, 1953

The Field Emitter: Fabrication, Electron Microscopy, and Electric Field Calculations*

W. P. Dyke, J. K. Trolan, W. W. Dolan, and George Barnes
Physics Department, Linfield College, McMinnville, Oregon

(Received November 7, 1952)

When a specially designed field cathode assembly was inserted in a commercial electron microscope which was appropriately modified for the purpose, the geometry of the needle shaped emitter was obtained from micrographs of several of its profiles at various stages of emitter fabrication and experimental use. An investigation of several methods of fabrication revealed that a smooth, simple, and relatively stable tungsten emitter geometry resulted from a refinement of the methods of Benjamin and Jenkins which combines the electrolytic etch and the smoothing action of surface migration. The electric field at any point on an emitter surface was calculated when the emitter geometry was precisely fitted with one equipotential surface from the family surrounding a charged, isolated sphere-on-orthogonal-cone. A theoretical surface distribution of current density for a typical emitter was derived from the calculated field distribution and the wave-mechanical field emission theory. From this result was calculated a value of the emitting area which was in good agreement with experiment.

INTRODUCTION

A STUDY of phenomena accompanying pulsed field emission at large current densities has been in progress in these laboratories.[1–3] The current densities predicted by the wave-mechanical image force corrected field emission theory[4] and obtained here experimentally are sufficiently greater than those reported previously[5] with direct current techniques to stimulate further research.

The interpretation of much of the field emission data that have been accumulated requires an accurate evaluation of both the electric field F and the work function ϕ at all points on the emitter surface, since according to the theory,[4] the current density J depends exponentially on the ratio $\phi^{\frac{1}{2}}/F$. Field emitters are conventionally needle shaped in order to provide the required value of the electric field, i.e., $10^7<F<10^8$ volts/cm, with convenient values of the applied potential. The emitter tip is usually a single crystal, approximately hemispherical with a radius of the order of 10^{-5} cm. Thus both ϕ and F vary appreciably over the emitting surface, the former by at least 10 percent in the case of tungsten[6,7] which is usually used, and the latter by as much as 30 percent as will be shown herein. The resulting variation in the ratio $\phi^{\frac{1}{2}}/F$ causes a total variation in J in excess of two orders of magnitude, illustrating the requirement for an accurate determination of the distributions of both ϕ and F.

The interpretation of the early field emission experiments, which are well summarized by Jenkins,[8] was limited by the inability of optical microscopy to resolve the emitter geometry upon which the calculation of both F and J depend. Haefer[5] improved the resolution of the geometry by use of electron micrographs of a single profile of each emitter. These revealed an emitter

* This work was supported by the U. S. Navy Office of Naval Research, the Microwave Laboratory of the University of California, and by Research Corporation.

[1] W. P. Dyke and J. K. Trolan, Phys. Rev. 82, 575 (1951).
[2] W. P. Dyke and J. K. Trolan, Phys. Rev. 85, 391 (1952).
[3] W. P. Dyke and J. K. Trolan, Phys. Rev. 85, 752 (1952).
[4] A. Sommerfeld and H. Bethe, *Handbuch der Physik* (J. Springer, Berlin, 1933), Vol. xxiv, No. 2, p. 441.
[5] R. H. Haefer, Z. Physik 116, 604 (1940).
[6] M. H. Nichols, Phys. Rev. 57, 297 (1940).
[7] G. F. Smith, Abstracts, Field Emission Seminar, Linfield College, McMinnville, Oregon (1952).
[8] R. O. Jenkins, Reports on Progress in Physics, the Physical Society (London) Vol. IX, 177 (1943).

configuration that was approximately conical with a hemispherical cap. Haefer idealized his emitters as hyperboloids of revolution[9] for the calculation of F_0 at the emitter apex, with considerable uncertainty caused by the dissimilarities between the experimental and the idealized geometries. Neither the hyperboloidal approximation nor the paraboloidal approximation introduced later by Becker[10] is useful for the required accurate calculation of the values of F at all points on the surface of the present emitters, as will be shown.

Such an evaluation of F is enhanced by a smooth, simple emitter geometry which can be sufficiently approximated by one member from a suitable family of equipotential surfaces. It is further desirable that the emitter geometry be relatively stable or at least well known during both the evacuation of the experimental tube, when the emitter is flashed at high temperature to remove surface contaminants, and during various experimental treatments to which the emitter may be exposed. Therefore an arrangement is required with which micrographs of several profiles of an emitter may be taken at various stages of its fabrication and use, permitting observations of the effects of those processes on the important emitter geometry. This was accomplished in the present work when a specially designed field cathode assembly was inserted in a commerical electron microscope which was modified for that purpose.

From an investigation of the geometries of emitters fabricated by several well-known and convenient methods, it was found that tungsten emitters of the desired geometries resulted from a modification of the process of Benjamin and Jenkins,[11] which combines the electrolytic etch with the subsequent smoothing action of surface migration during high temperature flash in vacuum.

The geometry of a present emitter was accurately fitted by one equipotential surface from the family surrounding a charged isolated sphere-on-orthogonal-cone permitting an accurate calculation of the electric field at any point on the emitter surface.

The present paper concerns field emitter fabrication, electron microscopy and surface electric field calculations; these results will be of further use in the interpretation of several other field emission experiments now in progress in these laboratories for which further reports are planned.[12] An accurate electric field distribution for the emitting surface will permit (1) a more precise comparison between experiment and the field emission theory,[4] (2) an evaluation of space charge effects. An accurate knowledge of the emitter geometry will be useful in the interpretation of mechanisms involved in the field emission initiated voltage breakdown.[3] For this purpose the emitter geometry is used in the calculation of both J and F just prior to breakdown. It is also required for the calculation of heat conduction and its generation by the resistive or the Nottingham[13] processes[13] which will be discussed in connection with voltage breakdown in another paper.

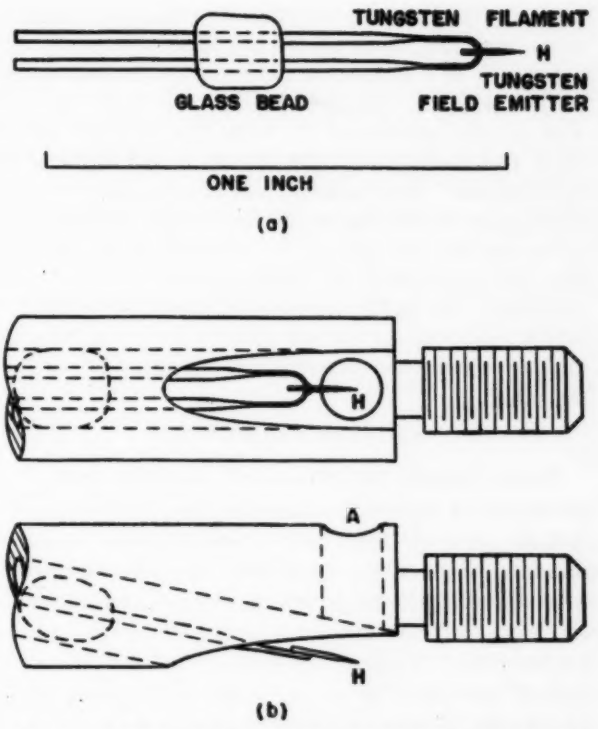


FIG. 1. (a) Field cathode assembly. (b) Modified sample rod for R.C.A. type EMT electron microscope with field cathode assembly inserted.

MICROSCOPY

The field emitter has usually been mounted on a hairpin filament which was heated to clean, outgas, and smooth the emitter by surface migration[14] in the final vacuum, the projection of the emitter beyond the filament being limited to about a millimeter in length in order to secure sufficiently high temperatures at the emitter tip. During the early stages of the present work, emitters were damaged when removed from the filament to obtain electron micrographs after use.† It was more difficult to remount the emitter on its filament for further experimental use as was sometimes required. Haefer[5] and Mueller[14] have also noted the difficulty with which electron micrographs of the field emitter were obtained.

The minute size of the emitter suggests a cathode

[9] Eyring, Mackeown, and Millikan, Phys. Rev. 31, 900 (1928).
[10] J. A. Becker, Bell System Tech. J. XXX, No. 4, Part 1, 907 (1951).
[11] M. Benjamin and R. O. Jenkins, Proc. Roy. Soc. (London) (A) 176, 262 (1940).
[12] W. P. Dyke and J. K. Trolan, Phys. Rev. 89, 799 (1953).
[13] W. B. Nottingham, Phys. Rev. 59, 907 (1941).
[14] E. W. Mueller, Z. Physik 126, 642 (1949).

† An R.C.A. type EMU electron microscope was made available to this project through the cooperation of F. B. Farquharson of the Engineering Experiment Station, University of Washington, Seattle.

assembly designed for insertion into the electron microscope as a unit, thus permitting repeated micrographs after each of a series of experiments. This was accomplished with the arrangement shown in Fig. 1A, which was readily positioned in a modified sample holder (Fig. 1B) acceptable in the R.C.A. type EMT electron microscope.[15] The emitter was mechanically rotated with a pair of tweezers to obtain micrographs of several of its profiles. Sample holder assemblies are shown in Fig. 2A, as supplied with the microscope, and Fig. 2B as modified. The latter was divided into two sections; *B* which remained in the instrument as a vacuum seal and *C* which was removed to accept the cathode assembly.

EMITTER FABRICATION

Needle shaped tungsten field emitters have been mechanically ground,[9] chemically etched[5,14] and electrolytically etched[11] from wire stock; emission has been obtained from the jagged end of a broken wire.[16] Mechanically ground emitters were judged from their electrical behavior to have sub-microscopic surface irregularities.[9] Electron micrographs of the chemically etched emitters[5] revealed undesirably complicated geometries. The electrolytic etch appeared more promising in view of the electrical behavior of field emitters fabricated in this manner by Benjamin and Jenkins[11] although electron micrographs of their emitters were not obtained. The electrolytic etch was used successfully in the present work.

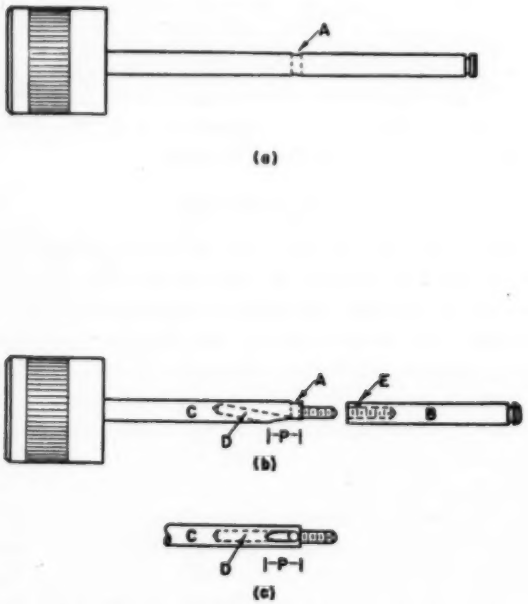


FIG. 2. Sample rods for R.C.A. type EMT electron microscope, (a) normal rod, (b) and (c) modified rod. Section B remains in microscope as vacuum seal while Sec. C is removed to insert field cathode assembly in cavity *D*. Electron beam at *A*.

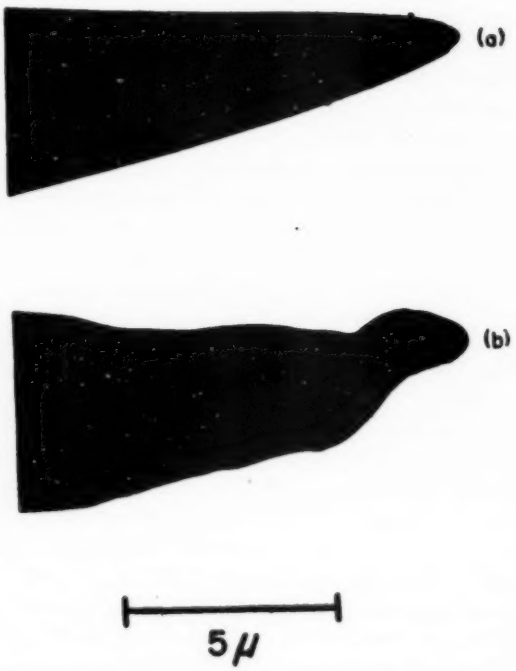


FIG. 3. Electron micrographs of a tungsten field emitter fabricated from Callite 200*H* 5 mil wire (a) following etch without prior heat treatment, (b) after subsequent heating at 2800°C for five minutes in vacuum.

The emitter blank, a short length of 5-mil tungsten wire (Callite 200H) was spotwelded in hydrogen to the apex of a support filament (Fig. 1A). The latter was formed from 20-mil tungsten wire which was beaded with uranium glass to insure mechanical rigidity. Its apex was reduced to 5-mil diameter and the entire assembly was then smoothed, both by electrolytic etch. Small projections were removed in the process in order to minimize undesired field emission. The emitter tip was then shaped by a final etch. For the several etches, an electrolyte of normal NaOH was used in which the emitter was one electrode and a nickel helix the other. A 60 cycle alternating potential of the order of 10 volts was used.

Electron micrographs of a typical emitter following etch (Fig. 3A) revealed a gross geometry which was approximately paraboloidal. The minute surface irregularities which are evident were correctly inferred by Benjamin and Jenkins from field emission patterns. These irregularities were smoothed by the surface migration of tungsten-on-tungsten,[14] and surface contaminants were simultaneously removed when the emitter was heated to 2200°C or more, in vacuum; however its gross geometry was undesirably altered as shown (Fig. 3B).

The process was repeated with a wire sample which had been preheated at 2800°C for 5 minutes in vacuum prior to the etch. Typical emitter geometries following

[15] J. H. Reisner and E. G. Dornfeld, J. Appl. Phys. **21**, 1131 (1950).

[16] Stern, Gossling, and Fowler, Proc. Roy. Soc. (London) (A) **124**, 699 (1929).

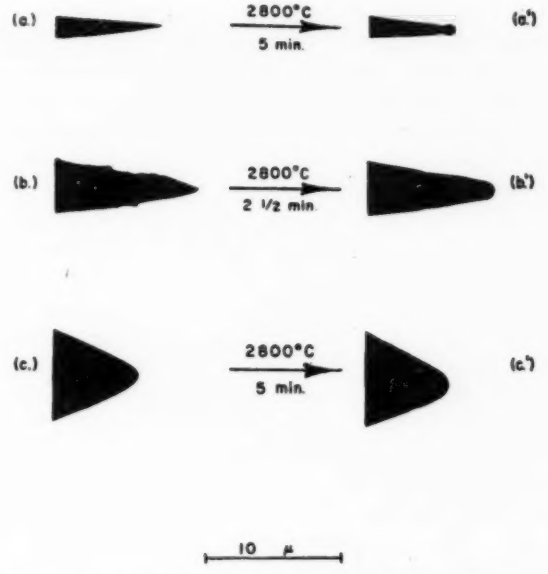


FIG. 4. Electron micrographs of tungsten field emitters electrolytically etched from 5 mil Callite 200*H* tungsten wire which was preheated at 2800°C for five minutes in vacuum. (*a*, *b*, and *c*) following etch; (*a'*, *b'*, and *c'*) after subsequent heating in vacuum as indicated.

etch are shown in Fig. 4 (*a*, *b*, and *c*) and in Fig. 4 (*a'*, *b'*, and *c'*) after further heat treatment as indicated in vacuum. Gross changes in geometry were avoided by this process.

The emitter geometries shown in Fig. 4 (*a'*, *b'*, and *c'*) were relatively stable to additional heating in vacuum at 2200°C which was required to remove surface contaminants such as oxygen during the evacuation of the experimental tube.

It is noteworthy that the final emitter radius depended on the cone angle of the emitter after the adequate heat treatment indicated in Figs. 4 (*a'*, *b'*, and *c'*). The cone angle was readily controlled by the duration of the electrolytic etch. If the etch was interrupted just prior to an appreciable decrease in length of the emitter blank, an emitter of small cone angle resulted (Fig. 4a). Continued etch resulted in the successive forms Figs. 4b and 4c, the latter occurring when most of the emitter blank lying beneath the surface of the electrolyte was removed by etch.

The emitters in Figs. 4 (*a'*, *b'*, and *c'*) had stable geometries suited to a number of experimental uses; their smooth surface of simple configuration was amenable to improved electric field calculations which are presented in the following section.

ELECTRIC POTENTIAL AND FIELD CALCULATIONS

Accurate electric field values at all points on an emitter surface may be calculated with conventional potential theory when the emitter is closely fitted with a suitable equipotential surface.

Emitters produced by the techniques used here have two geometric characteristics in common with a sphere mounted on an orthogonal cone; (1) the emitter tip closely approximates a sphere at polar angles up to 30° or more relative to the apex, and (2) the emitter shank is essentially conical at distances from the apex large compared with the radius of the tip (Fig. 5). That these emitters are well approximated by equipotential surfaces belonging to the family which surrounds the charged, isolated sphere-on-orthogonal-cone is evident from Fig. 5. When one such surface replaces the "core" (i.e., the conducting sphere-on-orthogonal-cone) as cathode, the anode being another equipotential surface from the same family, the electric field at the cathode is readily calculable.

For the potential distribution surrounding a charged, isolated sphere-on-orthogonal-cone, one uses:[17]

$$V=[V_R/R^n][r^n-a^{2n+1}r^{-n-1}]P_n(\cos\theta) \qquad (1)$$

where r and θ are ordinary plane polar coordinates referred to the center of the sphere as origin and the pole of the sphere opposite its contact with the cone as zero direction; R is the anode distance, V_R the applied potential, a the radius of the sphere, and P_n the Legendre function with n chosen so that the function vanishes for $\theta=\alpha$, the exterior half-angle of the cone.

The theoretical anode corresponding to the equipotential system defined by Eq. (1) can be described by setting $V=V_R$. In this case the second term in the parentheses becomes negligible, and the polar equation

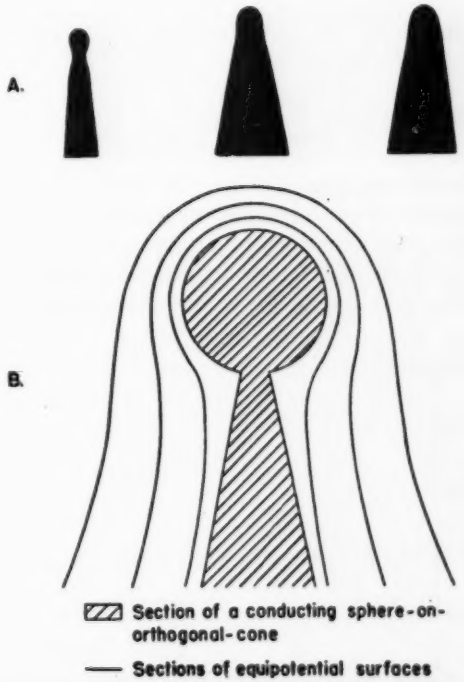


FIG. 5. A comparison between (*A*) typical field emitter geometries and (*B*) equipotential surfaces surrounding a charged isolated sphere-on-orthogonal-cone.

[17] W. R. Smythe, *Static & Dynamic Electricity* (McGraw-Hill Book Company, Inc., New York, 1949).

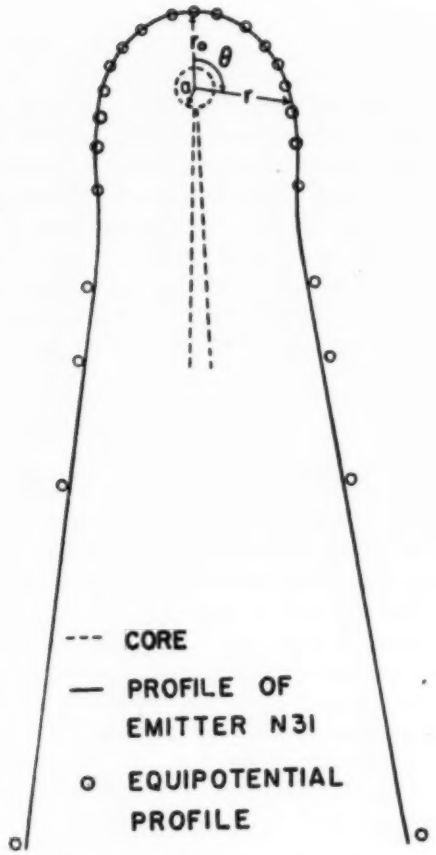


FIG. 6. An enlarged profile of emitter N-31 (from an electron micrograph) fitted with an equipotential surface from equation (1) with $n=0.10$, $a=1.235\times10^{-5}$ cm, and $r_0=4.00\times10^{-5}$ cm.

of the anode is:

$$r^n=R^n/P_n(\cos\theta). \qquad (2)$$

The surface described by Eq. (2) somewhat resembles a paraboloid. It can be readily approximated in the construction of an experimental anode.

Study of the characteristics of such a family of surfaces (Eq. (1)) reveals three important properties: (1) the neck constriction decreases and disappears as the potential above that of the cathode increases; (2) each surface approximates a sphere near the vertex; however at intermediate distances from the vertex it is approximately conical with larger interior angle than that of the core; (3) in the limit at very large distances from the apex the surfaces are parallel to the core.

To draw a chosen equipotential approximating a given emitter, one has three variables in Eq. (1) at his disposal: n, related through the Legendre function to the cone angle of the core, a, the radius of the core, and V, the potential above that of the core. After some experimentation it is possible to find a combination of these parameters that leads to a surface closely approximating a given emitter profile.

The equipotential surface fitted to emitter N-31, as shown in Fig. 6, was derived from Eq. (1) where $n=0.10$, $a=1.235\times10^{-5}$ cm and $r_0=4.00\times10^{-5}$ cm

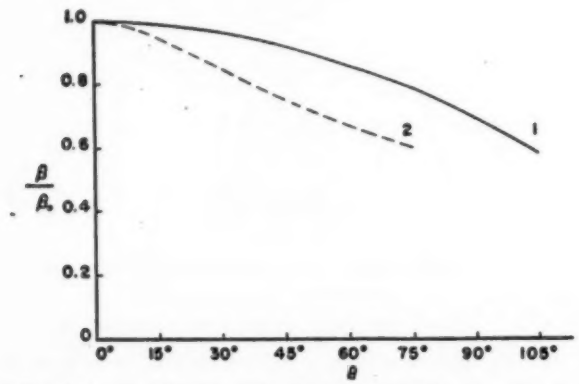


FIG. 7. A graph showing the dependence of the ratio β/β_0 upon polar angle θ for emitter N-31. Curve 1 from equation (4) with $n=0.10$, $a=1.235\times10^{-5}$ cm, and $r_0=4.00\times10^{-5}$ cm; Curve 2 from the paraboloidal approximation (Reference 10).

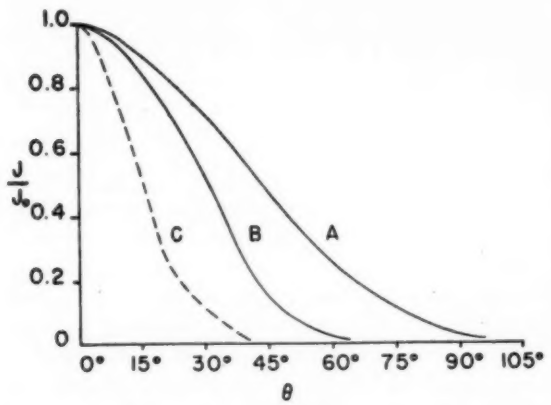


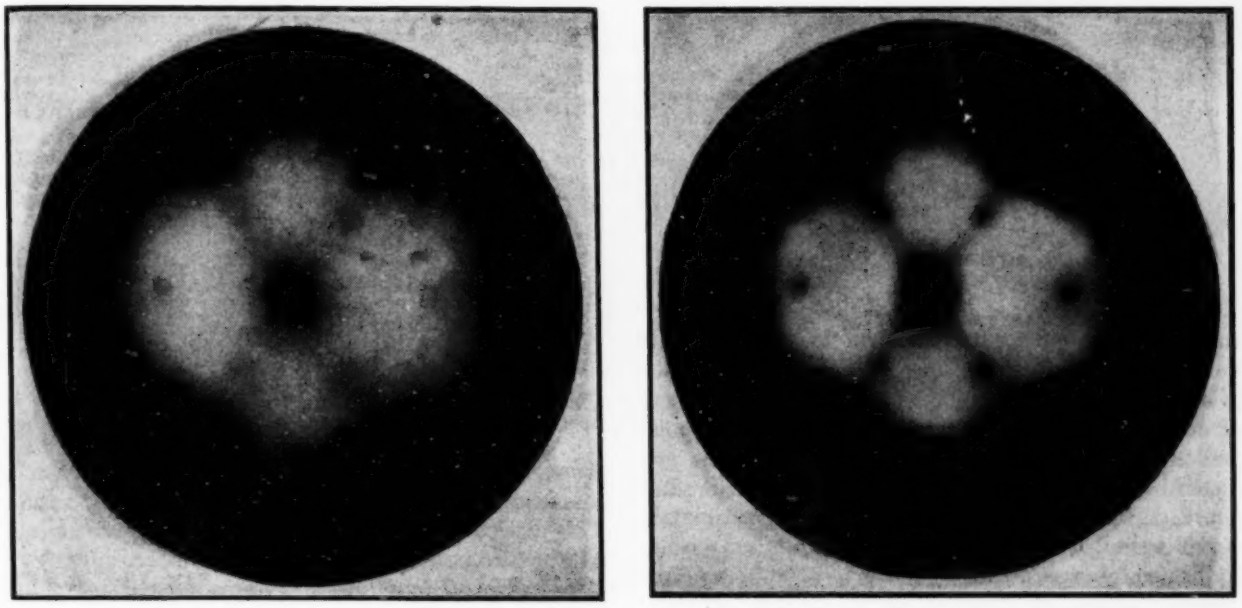
FIG. 8. A graph showing the dependence of the ratio J/J_0 upon polar angle θ for emitter N-31 calculated from the field emission theory (reference 4) with $\phi=4.5$ ev. Curves A and B resulted from the data of Curve 1, Fig. 7 with $F_0=7\times10^7$ volts/cm and $F_0=3\times10^7$ volts/cm, respectively. Curve C is from the data of Curve 2, Fig. 7 with $F_0=5\times10^7$ volts/cm.

(r_0 is the value of r at $\theta=0°$ on the surface which approximates the emitter).

The quantity β, defined at any point as the ratio of the electric field to the difference in potential between the anode and the point in question, is a useful measure of emitter geometry, being independent of the numerical value of the potential. For the surface of Fig. 6, the value β_0 at the emitter apex is

$$\beta_0=\frac{F_0}{V_R-V}=\frac{r_0^{n-1}}{R^n}\left[n+(n+1)\left(\frac{a}{r_0}\right)^{2n+1}\right]\cdot\frac{V_R}{(V_R-V)} \qquad (3)$$

where V_R is the potential difference between core and anode, V is the potential on the surface which was substituted for the emitter, and V_R-V is the potential difference between the emitter (i.e., the substitute surface) and anode. F_0 is the electric field at the emitter apex, i.e., the gradient of V evaluated at $r=r_0$, $\theta=0°$.



A B

FIG. 9. Photographic reproductions (positives) of the emission pattern from emitter N-31 as viewed on the metal backed phosphor anode of a Mueller type field emission projection tube. (A) $F_0=8.6\times10^7$ volts/cm, (B) $F_0=3.2\times10^7$ volts/cm.

Thus for emitter N-31 $\beta_0=3.6\times10^3$ cm^{-1}, which is the best available value. The present strong dependence of β_0 upon cone angle (through the variable n) was overlooked when the hyperboloidal or paraboloidal approximations were applied to such emitters.

The values of β defined as $F/(V_R-V)$, at various points on the substitute cathode surface are found in the usual way when the electric field components are derived from the potential gradient, with the aid of Eq. (1). There results:

$$\frac{\beta}{\beta_0}=\left(\frac{r}{r_0}\right)^{n-1}\frac{\left\{\left[n+(n+1)\left(\frac{a}{r}\right)^{2n+1}\right]^2[P_n(\cos\theta)]^2+\left[1-\left(\frac{a}{r}\right)^{2n+1}\right]^2\left[\frac{d}{d\theta}P_n(\cos\theta)^2\right]\right\}^{\frac{1}{2}}}{\left[n+(n+1)\left(\frac{a}{r_0}\right)^{2n+1}\right]}. \tag{4}$$

The ratios β/β_0 which were thus calculated for emitter N-31 are graphed in Fig. 7, curve 1; values of β/β_0 that were predicted when the paraboloidal approximation[10] was applied to the same emitter are shown in curve 2. The data from curve 1 lead to a predicted emitting area which is in good agreement with experiment, as will be shown next; that in curve 2 was less useful.

When the values of β/β_0 from curve 1, Fig. 7 were used with the field emission theory[4] and a constant work function $\phi=4.5$ (for clean tungsten), there resulted the predicted dependence of current density J on polar angle θ shown in Fig. 8, curves A and B, for values of the electric field at the apex $F_0=7\times10^7$ volts/cm, $F_0=3\times10^7$ volts/cm, respectively. The experimental current density distribution is somewhat more complicated than those shown in Fig. 8 because of the presently ignored but known variation of ϕ with crystallographic direction.[6,7] Nevertheless the present treatment is adequate in order to calculate the average polar angle at which the boundary of the significant emission is found, leading to an accurate evaluation of the total emitting area.

Emission patterns, Figs. 9A and 9B, were obtained when emitter N-31 was operated in a Mueller projection

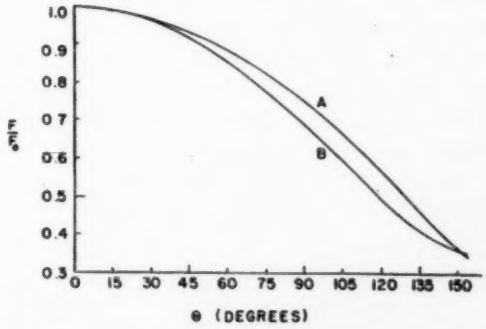


FIG. 10. A graph showing the dependence of the ratio F/F_0 upon polar angle θ for the typical emitters illustrated in Fig. 11. For Curve A, $n=0.14$, $r_0=2.1\times10^{-5}$ cm, $a=1.2\times10^{-5}$ cm. For Curve B, $n=0.17$, $r_0=1.5\times10^{-5}$ cm, $a=0.35\times10^{-5}$ cm. Curves A and B correspond to emitter profiles A and B of Fig. 11.

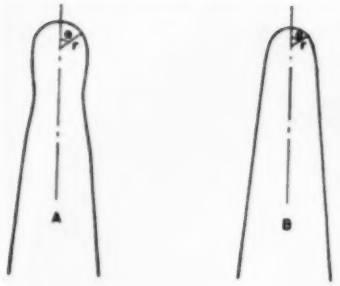


FIG. 11. Typical emitter profiles, (A) with extreme neck constriction, and (B) without constriction, corresponding, respectively, to curves *A* and *B* of Fig. 10.

tube[18] with $F_0=8.6\times10^7$ volts/cm and $F_0=3.2\times10^7$ volts/cm, respectively, the former when a current pulse of one microsecond duration was drawn. The periphery of emission occurs in the two cases at average polar angles that are in expected agreement with those predicted by curves *A* and *B*, Fig. 8. Curve *C* of the same figure, derived by use of the paraboloidal approximation (from curve 2, Fig. 7, with $F_0=5\times10^7$ volts/cm), is not confirmed by experiment.

Polar angles at various points on the emission patterns were identified from the known variation of current density with crystallographic direction;[18] the periphery of emission was located by a densitometric analysis of the calibrated photographic negatives of which Figs. 9A and 9B are positive prints. The patterns are approximately orthographic projections of the cathode surface emission.

The decrease of field with polar angle θ is further illustrated in Figs. 10 and 11, where the ratio F/F_0 is graphed for two widely different emitter profiles, with and without the frequently observed neck constriction. As one would expect, the field at the surface of the constricted emitter falls off less rapidly for small angles than in the case of the other emitter, but the fields differ nowhere by more than about 10 percent, and are again close together at values of θ corresponding to the constricted region (i.e., roughly $\theta=150°$). It is significant that at this remote part of the emitter the calculated field is still about 30 percent of its maximum value, a strong enough field to permit appreciable emission in this area if F_0 is near its highest experimental values. This expectation was borne out by the experimentally obtained emission patterns for such cases.

[18] E. W. Mueller, Z. Physik **106**, 541 (1937).

TABLE I. Values of $P_n(\cos\theta)$ for various values of n and θ.

θ	$n=0.21$ $\alpha\simeq169°$	$n=0.14$	$n=0.11$	$n=0.01$	θ	$n=0.16$ $\alpha\simeq175°$
15°	0.996	0.997	0.998	0.9998	13°	0.997
30°	0.982	0.989	0.992	0.9993	29°	0.986
45°	0.960	0.975	0.981	0.9984	39°	0.975
60°	0.928	0.955	0.965	0.997	55°	0.950
75°	0.886	0.927	0.944	0.995	65°	0.929
90°	0.831	0.892	0.917	0.993	75°	0.904
105°	0.762	0.850	0.882	0.990	90°	0.872
120°	0.673	0.789	0.837	0.986	100°	0.842
135°	0.557	0.712	0.776	0.981	110°	0.803
150°	0.395	0.602	0.692	0.974	120°	0.756
165°	0.250	0.420	0.565	0.959	130°	0.701
					140°	0.632
					150°	0.543
					160°	0.450

Values of $P_n(\cos\theta)$ were obtained from a straightforward though laborious numerical evaluation of its definitive hypergeometric series[17] which in some cases was carried to 120 terms. Values of $P_n(\cos\theta)$ are given in Table I.

Further work is in progress to apply these results to a refined test of the field emission theory and to other experiments mentioned in the introduction. Haefer[4] found excellent agreement between experiment and theory when β_0 was calculated from the hyperboloidal approximation. Although that method when applied to the present emitters yields values of β_0 that are somewhat lower than those obtained from the method described above, it is not necessary to conclude at this point that the theory is correspondingly incorrect since the effect of the filament structure, previously ignored, is such as to compensate in part at least for the error introduced by use of the hyperboloidal approximation.

We gratefully acknowledge the considerable assistance rendered by J. L. Boling with electron microscopy and by students with several of the computations. L. M. Perry and D. M. Barney provided valuable technical assistance.

JOURNAL OF APPLIED PHYSICS VOLUME 24, NUMBER 5 MAY, 1953

Elasto-Plastic Stress-Optical Effect in Silver Chloride Single Crystals

L. E. GOODMAN AND J. G. SUTHERLAND
Department of Civil Engineering, University of Illinois, Urbana, Illinois
(Received November 28, 1952)

The use of silver chloride as a material for "photoelastic" stress analysis offers the possibilities of studying both elastic and plastic states of stress in a crystalline material on either a micro- or macroscale. It is necessary, however, to relate quantitatively the stress state and the observed relative retardation and extinction angle. In this paper these relationships are developed from a general theory of stress birefringence, according to a stress-dependent hypothesis. This hypothesis and the resulting analytical relationships are experimentally vindicated by measurements made on a variety of single crystal specimens of silver chloride tested in simple tension in the elastic and plastic stress ranges.

INTRODUCTION

SILVER chloride is an elasto-plastic material with many of the properties of metals of structural importance.[1] In particular, its behavior at room temperature is elastic at low stress levels, becoming ductile as the stress is raised. In addition, silver chloride is transparent and, being a cubic crystal, it is optically isotropic in the unstressed state. When stressed, it becomes doubly refracting so that its effect on polarized light, properly interpreted, can be used to measure the state of stress. The present work has been directed toward providing a quantitative relationship between the quantities which can be observed in the laboratory (relative retardation Δ and extinction angle φ are the ones most easily measured) and the state of elastic or plastic stress in a specimen.

Crystalline silver chloride is photoelastically anisotropic; the principal axes of the indicatrix do not correspond, in general, to the directions of principal stress. The angles between these two sets of directions depend on the relative magnitudes of the principal stresses and on the relative directions of the incident wave normal, the crystal axes and the stress axes. The retardation of the transmitted light also depends on these quantities. An analysis relating stresses—which need not be "elastic—," orientation of the crystal axes and relative retardation has been prepared.

Experiments have been performed to check these analytical relationships. Briefly, the experiments have consisted in subjecting a single-crystal specimen, the orientation of whose crystal axes had been determined by back-reflection x-ray photographs, to a known state of stress. The relative retardation and extinction angle have then been measured. From these measurements, the so-called stress-optical coefficients which appear in the theory can be determined. If the theory is correct, these quantities should be constant for a wide variety of crystal orientations and throughout the elastic and plastic stress range. This has been observed to be the case, at least within the limits of experimental error.

It is of some interest to speculate on the reasons underlying the stress-dependence of the photoelastic effect in a crystalline metal-like material. That the photoelastic effect should depend on stress rather than strain in this case is not unreasonable. Mueller[2] has indicated three sources of birefringence in stressed ionic cubic crystals: the anisotropy of the Lorentz-Lorenz force; the anisotropy of the Coulomb forces; and the deformation of the crystal lattice which occurs when load is applied to the specimen. According to Mueller, this distortion of the lattice changes the energy levels and transition probabilities of the optical electrons, in this way altering the refraction of the atoms. Therefore, it produces a preferred orientation for polarized light, with accompanying retardation and other optical effects. As long as the behavior of the specimen is elastic, these optical effects will be in proportion to the load; this in turn will be proportional to either (macroscopic) stress or strain. Plastic deformation takes place when whole packets of molecules glide relative to one another. Since this does not imply any further stretching of the lattice, but only a relative motion of one lattice unit with respect to its neighbor, it may be inferred that in the plastic range the photoelastic effect in a crystalline (or polycrystalline) material will be proportional to the stress rather than to the strain.

The distortion of the crystal lattice suggested by Mueller as a source of the photoelastic effect, together with a relatively simple glide mechanism, appears to play a dominating role in the stress-optical behavior of ductile crystals such as silver chloride. This observation may be supported by contrasting the linearity of the stress-optical effect we have observed in AgCl with the nonlinear results reported by Maris[3] for the essentially nonductile materials NaCl, KCl, CaF_2. Although NaCl and AgCl have the same crystal structure and their ions have the same electrical charge and approximately the same size, their macroscopic mechanical properties are quite different. This difference in mechanical properties is reflected in the photoelastic effect.

[1] Moeller, Schonfeld, Tipton, and Waber, Trans. Am. Soc. Metals **43**, 36–69 (1951). A. B. Stepanov, J. Tech. Phys. (U.S.S.R.) **19**, 205–217 (1949).

[2] H. Mueller, Phys. Rev. **47**, 947–957 (1935).

[3] H. B. Maris, J. Opt. Soc. Am. **15**, 194–200 (1927).

STRESS-OPTICAL EFFECT IN CUBIC CRYSTALS

"What may be called the elementary theory of double refraction caused by stress or strain, on the lines developed by F. E. Neumann, has been extended to natural crystals by F. Pockels in a series of memoirs[4] published in 1889 and 1890."[5] In these papers, Pockels works with the relations between double refraction and strain. A complete treatment of the problem is given by Pockels in his *Lehrbuch der Kristalloptik*.[6] In *A Treatise on Photo-Elasticity*,[5] E. G. Coker and L. N. Filon present the theory in terms of the relations between double refraction and stress. No essential difference comes into question in their investigations since perfect elasticity is assumed throughout.

Now let

$$OX,\ OY,\ OZ$$

denote the principal axes of a cubic crystal; if a_0 is the reciprocal of the refractive index, then the equation of the index ellipsoid† in the unstressed crystal is

$$a_0^2(X^2+Y^2+Z^2)=1. \tag{1}$$

That is, the indicatrix in an unstressed cubic crystal is a sphere. Let the components of the applied stress, which may or may not be an "elastic" stress, referred to the three crystal axes be $\sigma_X, \sigma_Y, \sigma_Z, \tau_{XY}, \tau_{XZ}, \tau_{YZ}$. The effect of these stresses is to distort the indicatrix into an ellipsoid of the form

$$B_{11}X^2+B_{22}Y^2+B_{33}Z^2+2B_{12}XY+2B_{31}ZX+2B_{23}YZ=1. \tag{2}$$

It is now assumed that the differences between the coefficients of this ellipsoid and the original sphere are small quantities of the first order, linear in the stresses. We have, in matrix notation,

$$\begin{bmatrix} B_{11}-a_0^2 \\ B_{22}-a_0^2 \\ B_{33}-a_0^2 \\ B_{23} \\ B_{31} \\ B_{12} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{21} & - & - & - & - & - \\ C_{31} & - & - & - & - & - \\ C_{41} & - & - & - & - & - \\ C_{51} & - & - & - & - & - \\ C_{61} & - & - & - & - & C_{66} \end{bmatrix} \cdot \begin{bmatrix} \sigma_X \\ \sigma_Y \\ \sigma_Z \\ \tau_{YZ} \\ \tau_{ZX} \\ \tau_{XY} \end{bmatrix}. \tag{3}$$

There are, therefore, 36 stress-optical coefficients‡ C_{ij} in the most general kind of crystal. In the present case, for cubic crystals, the axes OX, OY, OZ are all equivalent. By symmetry we therefore have relations of the types, $C_{66}=C_{44}$, $C_{33}=C_{11}$, $C_{12}=C_{13}=C_{31}$. Only three independent stress-optical coefficients remain and the above equations take on the form,

$$\begin{aligned} B_{11}&=a_0^2+C_{12}(\sigma_X+\sigma_Y+\sigma_Z)+(C_{11}-C_{12})\sigma_X \\ B_{22}&=a_0^2+C_{12}(\sigma_X+\sigma_Y+\sigma_Z)+(C_{11}-C_{12})\sigma_Y \\ B_{33}&=a_0^2+C_{12}(\sigma_X+\sigma_Y+\sigma_Z)+(C_{11}-C_{12})\sigma_Z \\ B_{23}&=C_{44}\tau_{YZ} \\ B_{31}&=C_{44}\tau_{ZX} \\ B_{12}&=C_{44}\tau_{XY}. \end{aligned} \tag{4}$$

In the case of an isotropic material, these equations must remain the same in form for any rotation of the coordinate axes; from this it follows that for an isotropic material, $C_{44}=C_{11}-C_{12}$.

In general, the stress components relative to the crystal axes are not known; however, these values can be readily determined if the stress components are known relative to any axes Ox, Oy, Oz. The transformation equations are,

$$\begin{aligned} \sigma_X&=\sigma_x l_1^2+\sigma_y m_1^2+\sigma_z n_1^2 \\ &\quad +2(\tau_{xy}l_1m_1+\tau_{xz}l_1n_1+\tau_{yz}m_1n_1) \\ \sigma_Y&=\sigma_x l_2^2+\sigma_y m_2^2+\sigma_z n_2^2 \\ &\quad +2(\tau_{xy}l_2m_2+\tau_{xz}l_2n_2+\tau_{yz}m_2n_2) \\ \sigma_Z&=\sigma_x l_3^2+\sigma_y m_3^2+\sigma_z n_3^2 \\ &\quad +2(\tau_{xy}l_3m_3+\tau_{xz}l_3n_3+\tau_{yz}m_3n_3) \\ \tau_{YZ}&=\sigma_x l_2l_3+\sigma_y m_2m_3+\sigma_z n_2n_3+\tau_{xy}(l_2m_3+m_2l_3) \\ &\quad +\tau_{xz}(l_2n_3+n_2l_3)+\tau_{yz}(n_2m_3+m_2n_3) \\ \tau_{XZ}&=\sigma_x l_1l_3+\sigma_y m_1m_3+\sigma_z n_1n_3+\tau_{xy}(l_1m_3+m_1l_3) \\ &\quad +\tau_{xz}(l_1n_3+n_1l_3)+\tau_{yz}(n_1m_3+m_1n_3) \\ \tau_{XY}&=\sigma_x l_1l_2+\sigma_y m_1m_2+\sigma_z n_1n_2+\tau_{xy}(l_1m_2+m_1l_2) \\ &\quad +\tau_{xz}(l_1n_2+n_1l_2)+\tau_{yz}(n_1m_2+m_1n_2) \end{aligned} \tag{5}$$

in which (l_1, m_1, n_1), (l_2, m_2, n_2), (l_3, m_3, n_3) are the direction cosines of OX, OY, OZ, respectively, relative to the Ox, Oy, Oz axes.

With these six stress components determined for any point in a crystal, and assuming that the stress-optical coefficients C_{11}, C_{12}, C_{44} are known, it is easy to obtain the equation of the index ellipsoid relative to the crystal axes: that is, substitute the B_{ij}'s determined by Eq. (4) into Eq. (2).

It is desirable to be able to calculate the refractive indices and axes of polarization for a wave traveling, say, parallel to Ox. To determine these factors, the magnitude and direction of the semimajor and semiminor axes of the trace of the index ellipsoid in the yz plane must be calculated. First the equation of the ellipsoid relative to Ox, Oy, Oz is obtained by means of the transformation

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} l_1 & m_1 & n_1 \\ l_2 & m_2 & n_2 \\ l_3 & m_3 & n_3 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix}. \tag{6}$$

Now setting x equal to zero, the equation of the elliptic

[4] F. Pockels, Wied. Ann. Physik Ser. 3, **37**, 144–172, 269–305, 372–388 (1889); **39**, 440–469 (1890).

[5] E. G. Coker and L. N. Filon, *A Treatise on Photo-Elasticity* (Cambridge University Press, Cambridge, 1931), p. 288.

[6] F. Pockels, *Lehrbuch der Kristalloptik* (B. G. Teubner, Leipzig und Berlin, 1906).

† The surface here designated as the "index ellipsoid" or "indicatrix" is referred to as the Fresnel ellipsoid by Coker and Filon (reference 5). Fresnel, however, deduced the wave surfaces from a single-surfaced ellipsoid having semi-axes equal to the velocities.

‡ These stress-optical coefficients are the same as the piezo-optical constants of Pockels (reference 6) save for a logical reversal of sign; i.e., $C_{ij}=-\pi_{ij}$.

section of the indicatrix in the yz plane is obtained. The principal semi-axes of this ellipse give the refractive indices and polarizing directions at the point.

If the coefficients of the index ellipsoid, when referred to Ox, Oy, Oz, are denoted by b_{11}, $b_{12}\cdots$ then the section normal to Ox is

$$b_{22}y^2+b_{33}z^2+2b_{23}yz=1. \tag{7}$$

Or, on setting $y=\rho\cos\varphi$, $z=\rho\sin\varphi$, then

$$1/\rho^2=b_{22}\cos^2\varphi+b_{33}\sin^2\varphi+2b_{23}\sin\varphi\cos\varphi. \tag{8}$$

To obtain the directions of the principal axes of this ellipse we set $\partial\rho/\partial\varphi=0$, therefore,

$$\tan 2\varphi=2b_{23}/(b_{22}-b_{33}). \tag{9}$$

The values of φ obtained from Eq. (9) are the angles between the planes of polarization and the Oy axis.

The magnitudes of the semiaxes of the trace of the indicatrix in the xy plane are obtained by substituting the value of φ, given by Eq. (9), into Eq. (8). Consequently, we obtain,

$$\rho_{\min}^{\max}=(2)^{\frac{1}{2}}\{(b_{22}+b_{33})\mp[(b_{22}-b_{33})^2+4b_{23}^2]^{\frac{1}{2}}\}^{-\frac{1}{2}}.$$

The refractive indices η_1 and η_2 of the two polarized waves passing through the crystal are equal to ρ max and ρ min, respectively. The square of the difference of the refractive indices is therefore

$$(\eta_1-\eta_2)^2=\frac{(b_{22}+b_{33})-\{(b_{22}+b_{33})^2-[(b_{22}-b_{33})^2+4b_{23}^2]\}^{\frac{1}{2}}}{(b_{22}b_{33}-b_{23}^2)}. \tag{10}$$

In an unstressed cubic crystal, the equation of the section of the index ellipsoid normal to Ox is $a_0^2(y^2+z^2)=1$, where $1/a_0=\eta=$the average index of refraction. When the crystal is stressed, this circle is deformed slightly to become the ellipse, $b_{22}y^2+b_{33}z^2+2b_{23}yz=1$. Since the major and minor axes of this ellipse are nearly equal to the radius of the original circle, the coefficient b_{23} and the difference of the coefficients b_{22} and b_{33} must be very small quantities. In view of this, Eq. (10) can be simplified by using the first two terms of the binominal expansion of the square root. Also, the denominator of this equation is very closely a_0^4 or $1/\eta^4$: Eq. (10) therefore, simplifies to

$$\eta_1-\eta_2=\frac{\eta^3}{2}[(b_{22}-b_{33})^2+4b_{23}^2]^{\frac{1}{2}}. \tag{11}$$

It should be noted that the approximations just made follow directly from the assumptions upon which Eqs. (3) are based.

If light is passed through a crystalline plate of thickness h, then the two resulting plane-polarized waves will have undergone absolute retardations Δ_1 and Δ_2 given by

$$\Delta_1=h(\eta_1-1),\qquad \Delta_2=h(\eta_2-1). \tag{12}$$

The relative retardation of the two polarized waves is therefore,

$$\Delta=(\Delta_1-\Delta_2)=h(\eta_1-\eta_2)$$

or, utilizing Eq. (11),[7]

$$\Delta=\tfrac{1}{2}\eta^3h[(b_{22}-b_{33})^2+4b_{23}^2]^{\frac{1}{2}}. \tag{13}$$

The relative phase difference, measured in radians, is $2\pi\Delta/\lambda$, λ being the wavelength in vacuum.

PLANE STRESS IN A CUBIC CRYSTAL PLATE

The experimental techniques of photoelasticity are best suited to the study of two-dimensional (plane stress) problems. It is therefore desirable to proceed, on the basis of the foregoing theory, with a more detailed study of the generalized plane stress problem.

Considering a single-crystal plate lying in the zy plane, then, in the case of plane stress,

$$\sigma_z=\tau_{xy}=\tau_{xz}=0. \tag{14}$$

Substituting these values into Eq. (5) and the resulting expressions for the stress components referred to the crystal axes into Eq. (4), the coefficients for the equation of the index ellipsoid are obtained in the form

$$\begin{aligned}
B_{11}&=a_0^2+C_{12}(\sigma_y+\sigma_z)+(C_{11}-C_{12})(\sigma_y m_1^2+\sigma_z n_1^2+2\tau_{yz}m_1n_1),\\
B_{22}&=a_0^2+C_{12}(\sigma_y+\sigma_z)+(C_{11}-C_{12})(\sigma_y m_2^2+\sigma_z n_2^2+2\tau_{yz}m_2n_2),\\
B_{33}&=a_0^2+C_{12}(\sigma_y+\sigma_z)+(C_{11}-C_{12})(\sigma_y m_3^2+\sigma_z n_3^2+2\tau_{yz}m_3n_3),\\
B_{23}&=C_{44}[\sigma_y m_2m_3+\sigma_z n_2n_3+\tau_{yz}(n_2m_3+m_2n_3)],\\
B_{31}&=C_{44}[\sigma_y m_1m_3+\sigma_z n_1n_3+\tau_{yz}(n_1m_3+m_1n_3)],\\
B_{12}&=C_{44}[\sigma_y m_1m_2+\sigma_z n_1n_2+\tau_{yz}(n_1m_2+m_1n_2)].
\end{aligned} \tag{15}$$

The equation of the index ellipsoid at a point in the stressed crystal is

$$B_{11}X^2+B_{22}Y^2+B_{33}Z^2+2B_{31}ZX+2B_{23}YZ+2B_{12}XY=1. \tag{16}$$

We require the equation for the trace of this ellipsoid in the yz plane, consequently, we set $x=0$ and substitute,

$$X=m_1y+n_1z,\qquad Y=m_2y+n_2z,\qquad Z=m_3y+n_3z. \tag{17}$$

Equation (16) then becomes

$$\begin{aligned}
&[B_{11}m_1^2+B_{22}m_2^2+B_{33}m_3^2+2B_{31}m_1m_3+2B_{23}m_2m_3+2B_{12}m_1m_2]y^2\\
&+[B_{11}n_1^2+B_{22}n_2^2+B_{33}n_3^2+2B_{31}n_1n_3+2B_{23}n_2n_3+2B_{12}n_1n_2]z^2\\
&+2[B_{11}m_1n_1+B_{22}m_2n_2+B_{33}m_3n_3+B_{31}(m_1n_3+n_1m_3)\\
&+B_{23}(m_2n_3+n_2m_3)+B_{12}(m_1n_2+n_2m_1)]yz=1.
\end{aligned} \tag{18}$$

[7] J. F. Nye, Part II, Proc. Roy. Soc. (London) A200, 47–66 (1950).

This is an equation of the form of Eq. (7); that is, of the form, $b_{22}y^2+b_{33}z^2+2b_{23}yz=1$. Using Eq. (15) for the B_{ij}'s and well-known relationships between the direction cosines involved, we obtain

$$\begin{aligned}
b_{22}&=a_0^2+C_{12}(\sigma_y+\sigma_z)+C_{44}\sigma_y\\
&\quad+[C_{11}-C_{12}-C_{44}][\sigma_y(m_1^4+m_2^4+m_3^4)\\
&\quad+\sigma_z(m_1^2n_1^2+m_2^2n_2^2+m_3^2n_3^2)\\
&\qquad+2\tau_{yz}(m_1^3n_1+m_2^3n_2+m_3^3n_3)],\\
b_{33}&=a_0^2+C_{12}(\sigma_y+\sigma_z)\\
&\quad+[C_{11}-C_{12}-C_{44}][(m_1^2n_1^2+m_2^2n_2^2+m_3^2n_3^2)\sigma_y\\
&\quad+\sigma_z(n_1^4+n_2^4+n_3^4)\\
&\qquad+2\tau_{yz}(m_1n_1^3+m_2n_2^3+m_3n_3^3)]+C_{44}\sigma_z,\\
b_{23}&=C_{44}\tau_{yz}+[C_{11}-C_{12}-C_{44}]\\
&\quad\times[\sigma_y(m_1^3n_1+m_2^3n_2+m_3^3n_3)\\
&\quad+\sigma_z(m_1n_1^3+m_2n_2^3+m_3n_3^3)\\
&\qquad+2\tau_{yz}(m_1^2n_1^2+m_2^2n_2^2+m_3^2n_3^2)].
\end{aligned}\tag{19}$$

These coefficients together with Eqs. (9) and (13) give the polarization (extinction) angle φ and the relative retardation Δ, respectively.

In the case of a pure tension stress directed along the y axis, the coefficients have the relatively simple form,

$$\begin{aligned}
(b_{22}-b_{33})&=\sigma_y[C_{44}+(C_{11}-C_{12}-C_{44})(m_1^4+m_2^4\\
&\qquad+m_3^4-m_1^2n_1^2-m_2^2n_2^2-m_3^2n_3^2)],\\
b_{23}&=\sigma_y(C_{11}-C_{12}-C_{44})(m_1^3n_1\\
&\qquad+m_2^3n_2+m_3^3n_3).
\end{aligned}\tag{20}$$

EXPERIMENTAL STUDY OF SINGLE CRYSTALS

By applying a known state of stress to a plane crystal of known orientation and by measuring the relative retardation of normally incident light waves, the foregoing theory can be checked experimentally. In the tests to be described, single crystal specimens were subjected to pure tension loads. The orientations of the crystal axes were determined by back-reflection x-ray photographs, and the relative retardation of a monochromatic light ray was measured by means of a Berek compensator. Measurements were made in both the elastic and plastic stress ranges.

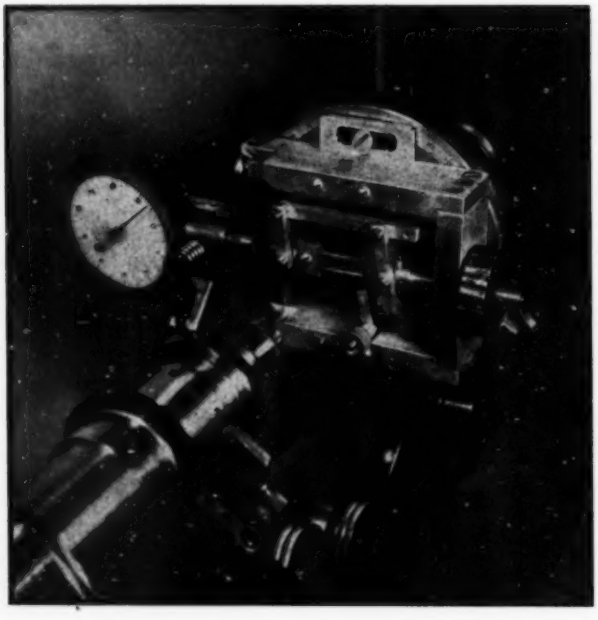


FIG. 1. Apparatus (on microscope stage) for applying tensile loads to specimens.

FIG. 2. Start of yielding in specimen *S*-10. Stress =440 psi. Nicols at −9° and +81° to the (horizontal) load direction. ×14.

Disks of clear silver chloride 3.25 in. in diameter and 0.3 of an in. thick were purchased from the Harshaw Chemical Company. These disks contained large single crystals which were detected by lightly etching the surface with a sodium thiosulfate solution. Single crystal slices about 0.1 in. thick were cut from the disks with a jeweler's saw. The specimens were then hand polished on metallurgical abrasive papers, utilizing kerosene as a lubricant. After washing the polished crystal in carbon tetrachloride, the surfaces were etched away by waving the specimen in a solution of sodium thiosulfate. The crystals were then rinsed in water and blotted dry with clean filter paper. The resulting specimens were about 0.04 of an in. thick, 0.28 of an in. wide, and 1.0 of an in. long.

The crystals were annealed using a technique described by J. R. Haynes.[8] Silver chloride has a high coefficient of thermal expansion and is extremely soft at elevated temperatures, consequently, care must be taken in the annealing process. The crystals were placed on a layer of finely ground quartz powder on top of a glass-topped table in the center of a small thick-walled copper box. In this box, the crystals were heated to

[8] J. R. Haynes, Rev. Sci. Instr. 19, 51 (1948).

FIG. 3. Glide bands in specimen *S*-10. Stress =490 psi. Nicols at −9 and +81 to the (horizontal) load direction. ×14.

740°F for about three hours after which they were very slowly cooled in the furnace.

Examination of the annealed specimens in polarized light revealed that, for all practical purposes, they were entirely free from stress. Some stresses were introduced while clamping the specimens in the loading device (see Fig. 1), however, these were very small; on examination between crossed Nicols the specimens initially appeared quite dark. When a crystal was subjected to a tensile load, such that the stresses were in the elastic range, considerable light was transmitted with almost uniform intensity over the whole specimen.

The onset of plastic action was indicated by either the growth of birefringent bands or of surface "hair-lines" or by both. The birefringent bands are caused by internal stresses remaining in the material after deformation by pencil glide. Generally, the birefringent bands are accompanied by surface lines; these are formed by the intersections of glide surfaces with the crystal surface. If the glide direction happens to be parallel to the surface, the lines will be practically invisible. On the other hand, surface lines will appear without birefringent bands if the glide surfaces overlap sufficiently.

Figure 2 shows the onset of plastic action in a typical specimen, *S*-10. In this case, there were two glide systems operating simultaneously; the angles between the bands and the stress axis were +38° and −56°, measured in the plane of the specimen. There were no apparent surface glide lines at this load, indicating that the glide directions were very nearly parallel to the surface of the crystal. In the present case, the angle between the two glide directions was approximately 94°. This is substantially in agreement with J. F. Nye's conclusion[9] that the glide direction in silver chloride is [110], since, in cubic crystals, the [110] directions make angles of either 60° or 90° with each other.

As the load was increased, other glide surfaces grew between those already present (see Fig. 3). At higher stresses, the glide surfaces generally became so closely spaced and overlapping that the specimen appeared almost homogeneous. Figure 4 shows a double system of surface lines observed in specimen *S*-22. This illustrates the development of surface lines without corresponding birefringent bands.

The orientations of the specimen crystal axes were determined by back-reflection Laue x-ray photographs. The photographs were interpreted by means of a Greninger chart together with a Wulff stereographic net and a standard projection for a cubic crystal. In some cases x-ray photographs were taken both before and after annealing the crystal; the heat treatment described did not affect the original orientation.

The cross-sectional dimensions of each annealed specimen were obtained using a micrometer caliper. Various reference lines were lightly scratched on the surface of the crystal, after which it was clamped in a small loading frame mounted on the stage of a Leitz polarizing microscope (Fig. 1). A measured tension load was then applied to the specimen and various measurements were made. The angles from the direction of the applied load to the planes of extinction and subtraction ($\varphi+45°$) were observed at various marked points. Using a Berek compensator, the relative retardations of

FIG. 4. Surface lines in specimen *S*-22. Stress =580 psi. Nicols horizontal and vertical, load applied horizontally. ×42.

[9] J. F. Nye, Part I, Proc. Roy. Soc. (London) **A198**, 190–204 (1949).

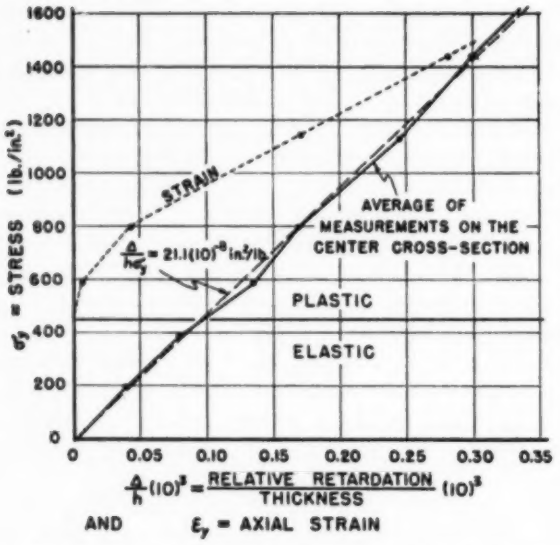


FIG. 5. Typical relative retardation *versus* stress curve, specimen *S*-10.

monochromatic (5461A) light waves were measured at these same points. In addition to these measurements, strains in both the axial and transverse directions were measured using a screw-micrometer eyepiece. Measurements were made for loadings corresponding to stress increments of about 200 to 400 lb per sq in.; the specimens were unloaded before each increase in load.

Stresses were computed based on the actual area of the strained specimen; this was determined by dividing the initial measured area by unity plus the measured axial strain. The observed retardations Δ were reduced to retardation per unit thickness, Δ/h, based on the actual thickness at the time of the observation. This thickness was determined by dividing the corrected area by the specimen width corrected in proportion to the measured transverse strain.

Eight groups of specimens were tested, each group

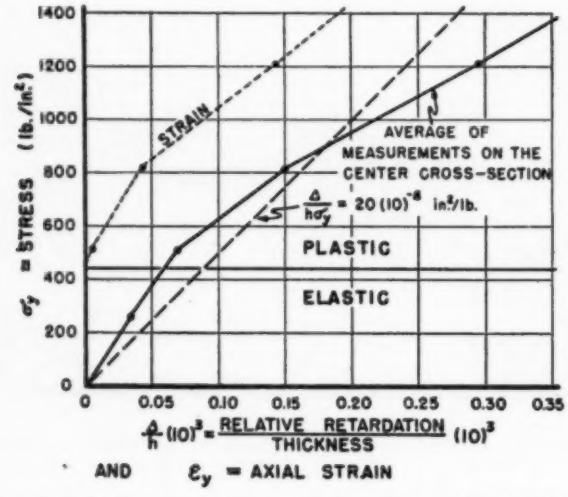


FIG. 6. Relative retardation *versus* stress for specimen *S*-9; this specimen showed the least linear stress optical effect.

having different orientations of the crystal axes relative to the principal stress and observation directions. For all except Group I, the specimens had one crystal axis (Oy) lying in the plane containing the axis of the specimen and normal to the surface of the specimen, that is, in the xy plane. The orientations of these specimens can therefore be described by two angles denoted

α = the angle between the Oz and the OZ axes or between the xy plane and the OX axis, and

Θ = the angle between the Oy and the OY axes (both of which lie in the xy plane).

In matrix notation, the direction cosines defining the crystal orientation are then,

$$\begin{bmatrix} l_1 & m_1 & n_1 \\ l_2 & m_2 & n_2 \\ l_3 & m_3 & n_3 \end{bmatrix} = \begin{bmatrix} \cos\alpha\cos\Theta & \cos\alpha\sin\Theta & -\sin\gamma \\ -\sin\Theta & \cos\Theta & 0 \\ \sin\alpha\cos\Theta & \sin\alpha\sin\Theta & \cos\alpha \end{bmatrix}. \quad (21)$$

In general, retardations were measured at three points on the central cross section of the specimen. The average of these three measurements (which, in general, were nearly equal) was then used to obtain a stress *versus* retardation graph such as shown in Fig. 5. The linearity of the stress-retardation curve shown in this figure is typical of those obtained for all specimens in Groups I, III, VII, and VIII, and also for specimens *S*-12 and *S*-13. The corresponding graphs for specimens of Groups II and VI were more erratic. In these specimens the retardation was difficult to measure

TABLE I. Summary of experimental results.

Group	Orientation α(deg)	Orientation Θ(deg)	Specimen No.	Ave. $\frac{\Delta}{h\sigma_y}$ (in.²/lb)	Ave. φ (deg)
I*	12.5	12	*S*-1	20.1 $(10)^{-8}$	73
			S-3	20.8 $(10)^{-8}$	73
			S-10	21.1 $(10)^{-8}$	72
II	57	30	*S*-4	9.8 $(10)^{-8}$	24
III	57	50.5	*S*-5	31.3 $(10)^{-8}$	12
			S-7	30.0 $(10)^{-8}$	14
IV	57	39	*S*-8	20 $(10)^{-8}$	11
			S-9	20 $(10)^{-8}$	8
			S-12	18.5 $(10)^{-8}$	12
V	22.5	0	*S*-13	29.8 $(10)^{-8}$	85
			S-14	28.7 $(10)^{-8}$	83
			S-22	29.0 $(10)^{-8}$	86
VI	22.5	25.5	*S*-15	6.7 $(10)^{-8}$	80∓
			S-16	5.7 $(10)^{-8}$	varied
VII	22.5	68	*S*-19	29.5 $(10)^{-8}$	32
			S-21	29.5 $(10)^{-8}$	33
VIII	22.5	89	*S*-17	31.8 $(10)^{-8}$	43

* Angles α and Θ given for Group I are only approximate. Calculations are based on the following measured orientation,

$$\begin{bmatrix} l_1 & m_1 & n_1 \\ l_2 & m_2 & n_2 \\ l_3 & m_3 & n_3 \end{bmatrix} = \begin{bmatrix} +0.959 & +0.187 & -0.215 \\ -0.212 & +0.973 & -0.099 \\ +0.191 & +0.141 & +0.971 \end{bmatrix}$$

TABLE II. Summary of analytical results.

Orientation as for group	Orientation α(deg)	Orientation θ(deg)	$\frac{\Delta}{h\sigma_y}$ (in.²/lb)	φ (deg)
I	12.5	12	23.4 $(10)^{-8}$	74.8
II	57	30	5.3 $(10)^{-8}$	17.5
III	57	50.5	30.9 $(10)^{-8}$	10.6
IV	57	39	18.8 $(10)^{-8}$	9.5
V	22.5	0	28.9 $(10)^{-8}$	90.0
VI	22.5	25.5	6.0 $(10)^{-8}$	77.1
VII	22.5	68	29.4 $(10)^{-8}$	31.6
VIII	22.5	89	33.0 $(10)^{-8}$	41.5

accurately, because it was quite small and was superposed on a double set of prominent birefringent bands. The stress-retardation curves for specimens *S*-14 and *S*-22 showed a slight concavity upward while those for specimens *S*-8 and *S*-9 were concave downward. Figure 6 (for *S*-9) shows the most nonlinear stress *versus* Δ/h relation obtained. It should be regarded as exceptional in the same way as Fig. 5 is typical.

In several cases, relative retardations were measured during reloading of a plastically strained specimen. In each case the retardation per unit thickness agreed quite well with that measured during the initial loading.

The results of all of the tests are summarized in Table I. The angles α and Θ define the crystal orientation of each group as determined from the x-ray photographs. The average retardations per unit thickness were obtained directly from graphs such as shown in Figs. 5 and 6.

COMPARISON OF EXPERIMENTAL AND THEORETICAL RESULTS

Before the theory discussed can be used to predict the optical effects in a stressed specimen, the stress-optical coefficients $(C_{11}-C_{12})$ and C_{44} must be determined. In the present study, these constants were determined such that the theoretical retardations would agree as closely as possible with the experimentally observed values. The accuracy of the theory is vindicated by the fact that these two constants could be chosen such that the optical effects predicted by the theory substantially agree with those observed for all of the specimens tested.

In Table II, the theoretical retardation coefficients and polarization angles are summarized for crystals oriented and stressed the same as the various groups of specimens tested. The values of the stress-optical coefficients used to compute the data given in this table are,

$$(C_{11}-C_{12})=-6.5(10)^{-8}\ \text{in.}^2/\text{lb}$$

and

$$C_{44}=+8.3(10)^{-8}\ \text{in.}^2/\text{lb}.$$

The refractive index, η, was taken as 2.07. The parameters Δ/h and φ were calculated using Eqs. (13) and (9), respectively, together with Eqs. (20). The dependence of these parameters on the crystal orientation is illustrated by Figs. 7 and 8. Also shown in these figures are the experimental results. Within the limits of experimental error, these agree with the theoretical predictions.

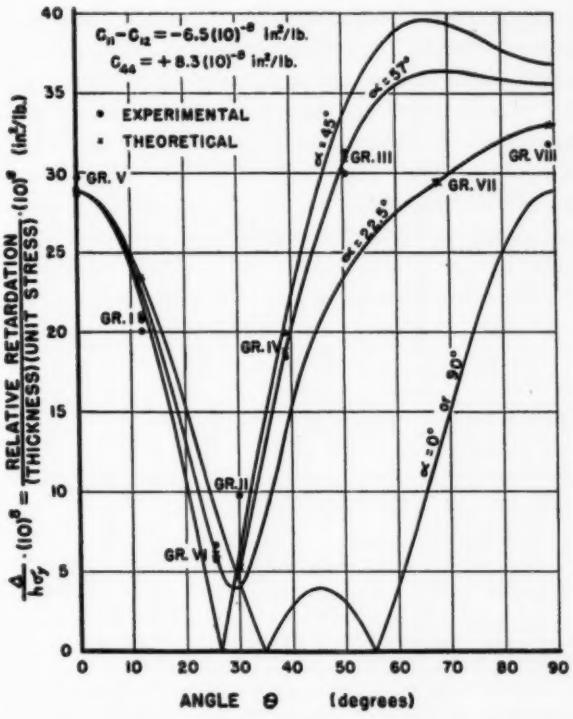


FIG. 7. Dependence of relative retardation on crystal orientation for uniaxial stress.

CONCLUSIONS

Tests of single-crystal specimens of silver chloride have shown that the optical effect in this crystalline

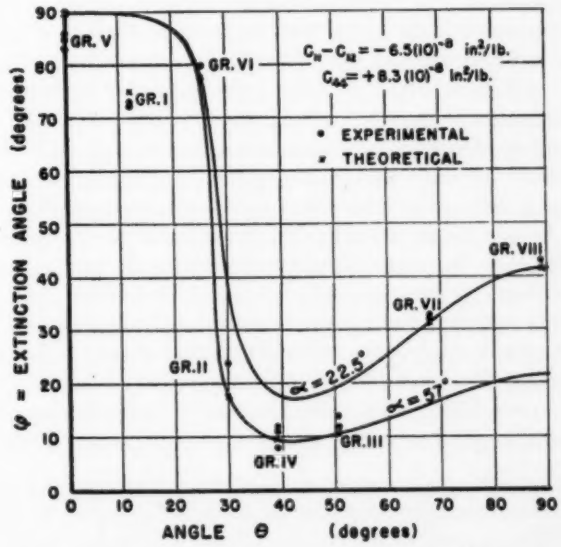


FIG. 8. Dependence of extinction angle on crystal orientation for uniaxial stress.

metal-like material depends on the stress state rather than on strain. The experimental data obtained from these tests have checked the accuracy, in both the elastic and plastic stress ranges, of a general theory which quantitatively relates optical effects to the stress state. In addition, these tests have provided data necessary for the determination of the so-called stress-optical coefficients (for silver chloride) which appear in the theory.

Silver chloride seems to be a very suitable material for "photoelastic" studies of the effects of plastic yielding on the state of stress in a crystalline material.

ACKNOWLEDGMENT

This work has been part of a study of the possibility of using optically active materials in the study of plastic states of stress, conducted at the University of Illinois, Department of Civil Engineering, under the sponsorship of the National Advisory Committee for Aeronautics. The authors are indebted to F. Seitz with whom some of the work has been discussed.

JOURNAL OF APPLIED PHYSICS VOLUME 24, NUMBER 5 MAY, 1953

The Gas Flow Past Slender Bodies

MAX M. MUNK
Naval Ordnance Laboratory, Silver Spring, Maryland
(Received November 7, 1952)

The following relates to moving objects immersed in a perfect gas otherwise at rest. The momentum of the resulting gas flow is defined. It is shown that it ultimately approaches a definite magnitude as the object assumes a steady motion, said magnitude being independent of the history of the previous motion.

This result is applied to the two-dimensional flows of airship theory, and thus the airship theory is extended to include compressibility effects to a limited extent.

INTRODUCTION

THE motion of a perfect fluid past slender bodies inclined by a small incidence angle relative to the direction of advance was studied many years ago by this author and forms the subject of his so-called airship theory.[1] The unsteady flow in air layers at right angles to the flight path, that is ordinarily in vertical air layers, stands in the center of the discussion (see Fig. 1). Use is made of the fact that this flow differs little from the plane two-dimensional flow for the boundary conditions in question so that the latter may be substituted for the respective components of the three-dimensional flow. Otherwise spoken, it is observed and utilized that the effectively incompressible air is flowing dominantly crosswise around and past the slender spindle and that the cross motion is only secondarily relieved by flow in the meridian planes.

As the flight velocities increased and as it became necessary to take the compressibility of the air into account, it was pointed out by Jones[2] that, and in what manner, the same thoughts and even the same equations apply also to supersonically moving slender bodies. The cross flow is still supposed to have small compressibility effects, the large and supersonic velocity being lengthwise in the steady flow picture and playing no decisive part in the unsteady cross flow. Missile bodies terminate ordinarily abruptly, flat bases being the rule. The cross flow produced immediately in front upstream the base persists then downstream the body. The cross flow is not suddenly annihilated at the base. The missile will, therefore, experience a resultant force equivalent to the effective momentum of the terminal cross flow.

The airship theory assumes gradual changes of the geometric dimensions along the missile axis. Fins contradicting this requirement call for special treatment.

APPARENT ADDITIONAL MASS

The airship theory takes account of the necessary information regarding the properties of the aerodynamic cross flow by utilizing the concept of the apparent additional mass of the cross-section contour. The fluid is supposed to be incompressible and free of viscosity, so that the ensuing flow has a potential and is uniquely determined by the instantaneous motion of the cross-section contour. Since at any point, positioned relative to the contour, the flow velocity v is exactly proportional to the velocity of advance V of the contour, the kinetic energy of the entire fluid motion set up by the contour moving in an otherwise resting fluid is proportional to the square of the velocity of advance. Hence it is equal to the kinetic energy of a fictitious mass moving with the same velocity as or being attached to the contour. This is the apparent additional mass. With a circular contour this mass is the same for motions of the contour

[1] Max M. Munk, "The Aerodynamic Forces of Airship Hulls," National Advisory Committee for Aeronautics Report No. 184 (1922).

[2] Robert T. Jones, "Properties of Low Aspect Ratio Pointed Wings at Speeds below and above the Speed of Sound," National Advisory Committee for Aeronautics Technical Note No. 1032.

in any direction. In the general case the kinetic energy associated with equal velocity of advance may be different for different velocity orientations and stands then in a more involved relation to the components of the velocity of advance.

The force acting on a missile slice per unit axis length is derived from the rate of change per unit time of the momentum m associated with the kinetic energy K

$$K=\bar{M}V^2\rho/2=\tfrac{1}{2}mV; \quad m=2K/V. \tag{1}$$

If the missile moves unsteadily itself these apparent momenta enter directly, the change being derived from the change of the velocity. If the body moves steadily itself at a small incidence angle, the rate of change of the momentum with reference to the axial coordinate enters.

The energy stored in the incompressible fluid is merely the kinetic energy of the fluid motion. Let ρ denote the constant fluid density, v the magnitude of the local velocity, and S the space volume. Then the kinetic energy is equal to the space integral

$$K=\int dSV^2\rho/2. \tag{2}$$

The integral (2) is finite, no infinite energy is required to set the fluid into motion. The integrand is essentially positive. The kinetic energy is concentrated in the vicinity of the moving body. When surrounding the body by a control or integration surface of increasing magnitude, concentric spheres say, and carrying out the integration successively through the volumes enclosed by larger and larger integration surfaces, larger and larger spheres say, the integral reaches asymptotically but very rapidly its upper and final value. The same does not depend on the selected shape of the control surfaces, provided the control surface increases its smallest distance from the body beyond any bound. Thus the kinetic energy (2) is well defined and is truly determinate. No special stipulation or limitation is necessary or in any manner called for.

An apparent additional momentum m can be computed from the kinetic energy K

$$m=2K/V. \tag{3}$$

It is not possible to compute this momentum immediately as a space-integral of the momenta of the individual fluid particles occupying the space elements dS. The momentum is a vector quantity. If for reason of symmetry the momentum is parallel to the velocity of advance, we may confine our attention to the momentum component parallel to the velocity of advance. The space integral

$$m=\int dSV\rho. \tag{4}$$

has then positive and negative contributions. The sum of all positive contributions by themselves is infinite

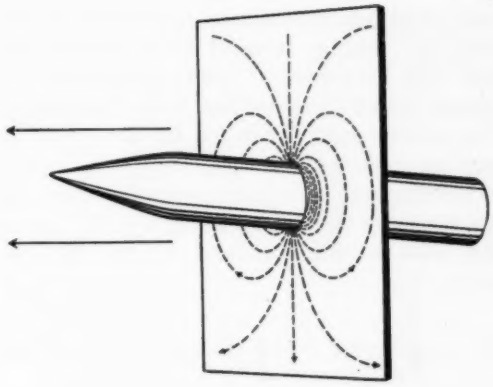


FIG. 1.

and so is the sum of all negative contributions by themselves. In consequence, the computed value of the integral depends entirely on the shape of the boundary or integration surface; however, this boundary may be enlarged and become distant from the body. It is thus seen that the space integral (4) without further stipulations does not define any definite momentum at all.

This is in harmony with the physical conditions. When setting up the flow, momentum is exchanged not only between the fluid surrounding the body and the body, but also between the fluid immediately surrounding the body and more outer portions of the fluid, surrounding said inner portions again. The fluid beyond the integration surface also contributes momentum, however distant that boundary may be from the body. An infinite amount of momentum changes place. The momentum contained in the entire fluid does not stand in any direct relation to the kinetic energy stored in the fluid which latter is equal to the energy transmitted from the body onto the fluid.

The term "momentum of the fluid" is made meaningful and definite, and for the present purpose the ambiguity and vagueness is removed by taking cognizance of the use we intend to make of the momentum. We are not interested in the momentum *per se* but we are only interested in the momentum as an intermediate quantity for the computation of the fluid force acting on the body. The momentum obtains its genuine meaning with this in view. The force on the body does not depend on the change of all momentum contained in the fluid or in any portion thereof, but it depends on and only on the momentum transmitted from the body to the fluid or vice versa. The term "momentum of the fluid" must and will therefore be interpreted as meaning that portion of the momentum, if a portion, which has been transmitted from body to fluid.

We may express the same thought by saying that the space integration for the computation of the momentum has to be carried out up to an outer integration boundary through which no momentum has been transmitted. It is sufficient to know that such boundaries exist; it is not necessary to determine them explicitly.

The process of integration may be carried out by first reducing the volume integral to a suitable surface integral. The existence of outer boundaries of zero-momentum transmittal justifies then disregarding the outer surface integral. Thus the integration is carried out over the surface of the body only.

The reduction to a surface integral is also used for the evaluation of the kinetic energy. With the incompressible fluid having constant density the essential integral is

$$2K/\rho = \int dS \nabla\phi \cdot V = \int do \cdot V\phi - \int dS\phi \nabla \cdot V, \quad (5)$$

ϕ denoting the potential of the velocity field, S denoting the volume, and o the boundary surface area. The right-hand space integral (5) is zero for the incompressible fluid. It is thus seen that the kinetic energy is substantially determined by the surface integral of the product of the flux and the potential. The portion of the pressure not balanced and not pertaining to the steady flow, that is the difference between the pressure and the Bernoulli pressure is equal to $\rho\partial\phi/\partial T$, T denoting the time. Hence (5) represents the time integral of the rate of energy hydraulically pumped or forced into the fluid while the flow is set up.

If the body executes a translationary motion V in constant direction the flux rate element is also equal to $doV \cdot do$. Hence the kinetic energy can also be derived from the surface integral

$$2K/V\rho = V \cdot \int do\phi. \quad (6)$$

Integral (6) is formally equal to the space integral for the momentum

$$m/\rho = \int dS \nabla\phi = \int do\phi. \quad (7)$$

The integration surface for zero-momentum transmittal is consistent with

$$\int do\phi = 0.$$

The flow momentum in the narrower and special sense explained above is, therefore, associated with the energy and is concentrated in the vicinity of the body like the energy. It is equal to ρ times the surface integral of the potential taken over the body surface.

INFLUENCE OF THE COMPRESSIBILITY. ENERGY

This paper is directed to the discussion of the compressibility in connection with the relations discussed. As the Mach number of the velocity of advance of the slender body, or as its angle of incidence increases, the Mach number of the cross flow will also increase and will become noticeably larger than zero. We assume the cross flow to remain subsonic but not incompressible.

We have accordingly to inquire into and to discuss the meaning of the energy and of the momentum of the three-dimensional flow set up by a body moving through a gas otherwise at rest. We assume the fluid to be free of viscosity and to expand barotropically, that is the density ρ is a function of the pressure p alone.

The ensuing gas flow is not solely determined by the instantaneous motion of the body. The flow depends in large measure on the past history of the motion of the body. The energy of the flow is also not determined by the instantaneous motion of the body as consequence thereof.

This energy consists of two distinct parts: the kinetic energy and the elastic energy. The kinetic energy is defined as with the incompressible fluid Eq. (2) taking, however, account of the variation of the density. The elastic energy represents the work stored in, or removed from, the fluid elements by virtue of their change of volume in conjunction with the pressure prevailing. The elastic work per mass element of the gas element is defined by

$$dW = \int pd\left(\frac{1}{\rho}\right), \quad (8)$$

and the volume element dS is occupied by the mass ρdS so that the elastic work per volume element becomes $\rho \int pd(1/\rho)$. Thus, the elastic work is basically equal to

$$W = \int^{S} dS\rho \int^{\rho} pd\left(\frac{1}{\rho}\right). \quad (9)$$

This integral (9) suffers from the same imperfections as the integral (7) for the momentum. The integrand assumes positive and negative values, the integral of the positive values only is infinite, and thus the integral has no definite value at all but the outcome of the integration depends on the integration process. The physical reason for this is also very similar. Elastic energy is not only exchanged between the body and the surrounding gas but also between an intermediary mass of gas and the outer mass of gas surrounding the intermediary mass. The reasoning and remedy removing the vagueness follows also the arguments laid down above. We are only interested in the elastic energy transmitted from body to the gas. At sufficiently large distances from the body the pressure becomes virtually constant, and equal to the ambient pressure. This ambient pressure delivers elastic energy equal to its magnitude multiplied by the change of volume of the enclosed gas. Hence the integration has to be performed for an outer control surface such that the volume of the enclosed gas is equal to the volume occupied by the same gas in the undisturbed state. It is again sufficient to realize that such integration surfaces exist. Again the volume integral is reduced to a suitable surface integral and only the body surface is considered in the integration.

In the case of a polytropically expanding gas with exponent γ this leads to the following expression for the elastic energy per unit volume

$$W = c_0^2\rho_0\left\{\left[\frac{(1-\frac{1}{2}(\gamma-1)M_0^2(V/V_0)^2)^{\gamma/(\gamma-1)}}{\gamma(\gamma-1)} - \frac{(1-\frac{1}{2}(\gamma-1)M_0^2(V/V_0)^2)^{1/(\gamma-1)}(1-\frac{1}{2}(\gamma-1)M_0^2)}{\gamma} + \frac{(1-\frac{1}{2}(\gamma-1)M_0^2)^{\gamma/(\gamma-1)}}{\gamma-1}\right]\right. \tag{10}$$

For very small density changes this reduces to $\frac{1}{2}c_0^2(\rho-\rho_0)^2/\rho_0^2$, a product containing the square of the density increment; c_0 denoting the ambient velocity of sound.

The total energy of the fluid is the sum of its kinetic energy and its elastic energy. It is considered that both be separately computed for any specific flow.

A simplification of the computation is possible if the Mach number of the steady flow is very small. Equation (5) becomes then approximately true and applicable. In (5) the potential but not the density has to be modified to reflect the somewhat larger Mach number. The steady almost incompressible flow is then supposed to decline or fade out at infinity almost like the incompressible flow. It is supposed that the motion is created infinitely gradually, with gradual change of the Mach number of the velocity of advance from zero to a small value. It is assumed that all these steady flows follow each other successively and continuously. Such a succession of steady flows is not really or strictly a solution of the equations of motion. For very small Mach numbers, however, the error becomes negligible. That is to say, the change of the potential is carried to terms proportional to the square of the Mach number, but the density is still considered equal to the ambient or undisturbed density. The total energy becomes thus approximately equal to

$$K=\tfrac{1}{2}\rho_0\int do\cdot v\phi', \tag{11}$$

ϕ' denoting the potential of the slightly compressible flow.

For the circular contour having the diameter $D=2R$, the most important contour in conjunction with the airship problem, the potential of the flow now two-dimensional is given by the expression[3]

$$\phi'=\phi_{M=0}+RVM^2\left[\left(\frac{13}{12}\frac{1}{r}-\frac{1}{2}\frac{1}{r^3}+\frac{1}{12}\frac{1}{r^5}\right)\cos\theta + \left(-\frac{1}{4}\frac{1}{r}+\frac{1}{12}\frac{1}{r^3}\right)\cos3\theta\right]+\cdots, \tag{12}$$

Rr and θ being the polar coordinates. This leads to the total energy

$$K=\tfrac{1}{8}V^2\rho_0D^2\pi[1+2M^2+\cdots] \tag{13}$$

in an expansion in terms of powers of the Mach number.

[3] Carl Kaplan, "Two Dimensional Subsonic Flows Past Elliptic Cylinders," National Advisory Committee for Aeronautics Report No. 624 (1938).

THE MOMENTUM OF THE GAS FLOW

We proceed now to the discussion of the momentum of the three-dimensional gas flow. The energy integral does not yield a definite apparent additional mass, because the energy not only fails to be proportional to the square of the velocity of advance but it is not even a unique function of the velocity of advance. In general, a body immersed in a compressible fluid does not move under the action of external forces as if its actual mass was increased by an apparent additional mass, not even a variable one depending on the velocity of advance. In consequence the energy does not lead to a corresponding momentum. The momentum cannot be computed from the energy but it is independent of it.

A study of the conditions leads now to the somewhat unexpected discovery that the momentum stands in a much simpler and more definite relation to the motion of the body than the energy does. To obtain clearer insight into this we have to analyze the ensuing motion in more detail. We assume that the fluid was first at rest. The body is next made to carry out a succession of motions finally assuming a constant velocity in constant direction. In the vicinity of the body, extending farther and farther out, the fluid will then assume a definite pattern which we will call quasi steady. It is not quite steady in that within finite time it never extends to infinity. It is rather bounded by a more or less spherical outer shell within which the fluid carries out an unsteady wave motion. The wave moves with velocity of sound towards infinity. The energy of the flow can be divided into two portions, the energy of the quasi-steady part and the wave energy. The energy of the quasi-steady flow assumes asymptotically, but rather rapidly, a definite value depending on the velocity of the body only. (The final velocity of advance of the body is assumed to be subsonic.) The energy carried away by the wave represents the additional energy depending upon the history of the body motion. In general it has no upper bound but depends entirely on how the motion was created. An upper bound exists under certain restrictions, namely, if the velocity of advance is always in definite direction and orientation and changes monotonously, so that the acceleration is

always in direction of and never opposite to the velocity of the body.

We have in mind a compressible fluid that expands barotropically, that is, the density is a function of the pressure alone. The fluid is free of viscosity and heat conduction. In that case there prevails a substantial difference between the behavior of the energy and that of the momentum in relation to the substantially developed steady motion. The energy transmitted from the body to the fluid depends on the mode of creation of the steady flow. The momentum does not. The same steady or quasi-steady flow is always associated with the same momentum transmitted from the body to the fluid, regardless of how the flow was created. This implies that if in a large and ever growing region about the body the fluid comes to rest again, because the body has come to rest and stays now at rest, no momentum has been transmitted *in toto* from the body to the fluid. The remaining wave system carries off energy but not momentum.

This relation is of significance in the two-dimensional case for extending the airship theory to cases of compressibility effects of the cross flow. The proposition has now to be demonstrated. We still have the three-dimensional flow in view. The absence of momentum of the actual wave system requires in the first place that really linear or acoustical sound wave systems, governed by the well-known equations of acoustics, are free of momentum. Acoustics disregards density differences in comparison with the basic density, and it disregards velocities in comparison with the velocity of sound. The velocity of sound is supposed to be constant. All this leads to linearization of the equations. Any wave motion can thus be obtained by linear superposition of symmetrical spherical waves and also by superposition by pairs of opposite plane waves. The superposition is linear, the momentum is also linear, and the elements of superposition are free of total momentum. It follows that any linear or acoustical wave is free of total momentum.

The actual wave is not precisely described by a linear equation. Indeed the separation of the flow into the wave and the quasi-steady flow portion takes place gradually and asymptotically only and is never quite completed. The two portions remain intermixed to some degree; the wave leaves the quasi-steady flow behind and deposits it, so to speak. The wave emerges more and more but never fully into a linear wave.

The wave occupies an ever increasing volume; it is contained between two concentric spheres of constant distance so that the wave volume increases proportional to the square of r, the distance from the body. The momentum per unit volume is not precisely proportional to the velocity, and hence to $1/r$, but it has also terms proportional to the square of the velocity, and higher terms. It has to be shown that all these quadratic terms cancel sufficiently so that the aggregate sum diminishes faster than proportional to $1/r^2$, at least like $1/r^3$.

Now the pertinent portion of the wave is bipolar, like a doublet. At the two ends of a diameter parallel to the direction of advance of a symmetric body, the velocity profile is sensibly the same in equal direction absolutely, but the one profile orientated inwardly the other outwardly. The momentum per unit wave front cancels locally as far as the first power of v is concerned. It is now noted that all momentum contributions ascribable to or porportional to the squares of the velocity are in the same directional relation to the radius, that is, they are all outwardly directed from center to periphery. This can for instance be seen with the density variation. If the velocity is outward, the density change is positive; if the velocity is inward, the density change is negative. In either case the momentum change has equal sign relative to the radial direction. And since this is the case with one such term, it must be the case with all. The square of the velocity has no direction. The radial direction is the only one available. There is no other sufficient reason why it should vary. Distinct discontinuities would follow from any other state of things. It would be impossible to write down the direction in any other way. Since now the profile is the same at opposite ends of the diameter only turned around, and since the quadratic terms have the same direction nevertheless, the remaining difference is ascribable to differences in the radial distances. These are by themselves of the order $1/r$ relative to each other. Hence the momentum will cancel up to terms proportional to $1/r^3$ will ultimately approach zero; whatever momentum the wave system carries is ultimately wholly and entirely deposited in the quasi-steady flow, as the passing wave builds it up.

It follows that the steady flow including any wave system associated with it is associated with a definite momentum depending on the body velocity only. This momentum can be expressed by the motion of an apparent additional mass, but this mass depends itself on the velocity of the body, and hence its use is not particularly useful.

COMPUTATION OF THE MOMENTUM

The momentum uniquely associated with the steady flow at a specific subsonic Mach number can be computed from the minimum energy associated with the successive steady flows. It is assumed that the flow is infinitely gradually produced by changing the Mach number M continuously. The infinite pressure changes giving rise to the change of momentum are not readily computed and are not even unique. The force can, however, be computed from energy considerations. The force necessary to accomplish the flow change is computed from the increase of the flow energy K. This

gives for the momentum m,

$$m=\frac{1}{c_0}\int_0^M \frac{dK}{dM}\frac{dM}{M}=\frac{1}{c_0}\int_0^M \frac{dK}{M}. \tag{14}$$

For small Mach numbers, the momentum is more conveniently computed from (11). For extremely small Mach numbers the momentum and minimum energy results the same as with the incompressible fluid.

The knowledge of the momentum enables us to compute the energy for a special but widely applicable case. Suppose the Mach number is very small but not zero. Suppose further the motion of the body is instantaneously created from zero and then kept constant, so that the moment is then transmitted to the fluid. The momentum is thus transmitted not with evenly increasing velocity but with constant velocity. The force f is an unknown function of the time, such that

$$\int f dT = K/V,$$

$$K_0 = mV/2, \tag{15}$$

$$K=\int V f dT = V\int f dT = mV = 2K_0, \tag{16}$$

f denotes the force required to act on the body, m denotes the momentum, V the velocity. It follows that by the described mode of creation of the flow the energy is doubled. Half of the energy stays in the quasi-steady flow and the other half is carried out by and into the wave portion of the flow.

For the special case of a sphere, the problem has been solved by means of full detail computations by Longhorn.[4] The special result obtained by him is in agreement with the general theory here presented.

The flow momenta occurring in the airship theory refer to two-dimensional flows caused by moving contours. Two-dimensional sound waves are almost analogous to three-dimensional ones; they are, however, not strictly contained between a pair of concentric circles moving approximately with the velocity of sound. They are that only substantially and effectively. The preceding proof relating to the three-dimensional case requires, therefore, some elaborations to make it applicable to the two-dimensional case. It is unnecessary to carry this out in the present paper as the formally two-dimensional flows of airship theory represent really aspects of three-dimensional flows about very elongated objects moving broadside.

[4] L. Longhorn, Quart. J. Mechanics and Appl. Math. V, Part 1 (1952).

APPLICATION TO THE AIRSHIP THEORY

The airship theory employs the momentum of the flow, not the energy of the flow. Furthermore, it assumes a very slender body or spindle, so that there is time for the steady flow to establish itself, provided the body is well inside the Mach cone originating at the nose of the airship. It appears that under these circumstances the airship theory is applicable, even if the cross flow is subject to moderate compressibility effects. The momentum of the cross flow has then to be computed for the particular Mach number of the cross flow in question. It has been shown in the present paper that this momentum does not depend on the mode of creation of the flow. Hence the momentum does not depend on the airship shape. The compressibility and the necessity to wait for the establishment of the steady flow will shift the cross forces slightly downstream.

The additional apparent mass increases with the Mach number. It follows that the lift slope must increase with the angle of incidence. It is true that the same increase is observed even in practically incompressible flow. But then the stern forms a sharp end, and viscosity and vorticity effects explain this; particularly at excessively large angles of attack. This explanation is less acceptable for spindles terminating in flat base ends. The curvature of the lift curve with such spindles is then to a dominant part explained by the change of the apparent additional mass of the cross contour or else by the increase of the momentum ascribable to the compressibility.

This author has analyzed a number of wind tunnel tests with this in view, using Eq. (13). The report presenting the outcome has not been published. The lift slope points computed from the tests scatter a great deal for a number of reasons. However, the computations based on the theory of this paper gives an order of magnitude of the lift slope variation in agreement on the average with the experimentally observed facts.

Space Charge Spread of Reflected Electron Beams Studied by a Photographic Method

J. T. WALLMARK
The Royal Institute of Technology, Stockholm, Sweden
(Received December 3, 1952)

The spread of electron beams of rectangular cross section is studied (a) in a field free space, (b) in a reflection space. An experimental method is tried to trace the electron paths from the excited gas molecules left behind by directly photographing the beam in a high vacuum. The results are checked with published theoretical calculations in case (a), and with results obtained through an iterative method in case (b).

INTRODUCTION

MANY attempts have been made to find convenient ways of studying electron paths and have resulted in successful practical methods such as the rubber model, electromechanical path tracers, graphical methods, etc., which are now widely used. However, when one wants to find the influence of space charge on the electron paths there are no accurate methods available and mathematical treatment is as yet confined only to special cases.

This article deals with the spread of an electron beam of rectangular cross section in a retarding electric field, a problem which arose in the development of a new kind of electron tube.[1] In order to study such a beam a photographic method has been tried that, although not new, has not been described earlier in spite of its obvious merits in handling complicated space charge problems.

Photographic methods have been employed for studying electron paths in a large number of cases since the discovery by Wehnelt[2] that electron paths could be made visible by keeping the rest gas pressure between certain limits. Then the space charge of the ions formed counter-balanced the negative space charge of the electrons and a well-focused "thread" ray resulted. Further work, especially that by Brüche,[3] showed that these thread rays are a vivid and accurate means of studying electron paths where space charge influence is not the chief concern.

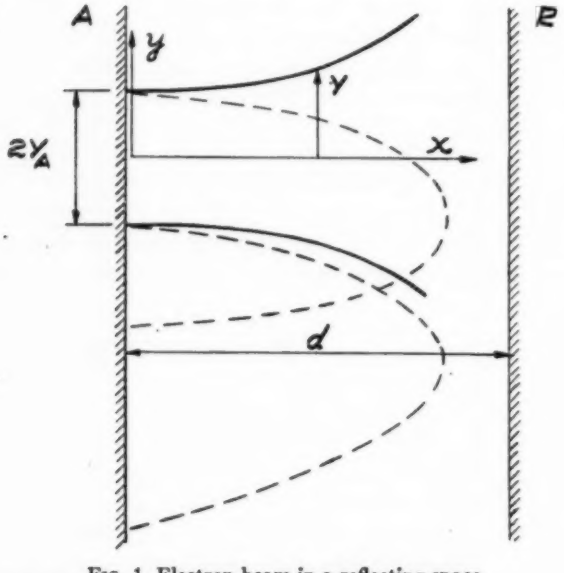


FIG. 1. Electron beam in a reflecting space.

If the rest gas pressure is reduced below a certain limit the space charge balance is lost and the electron beam loses its character of thread ray, spreads, and soon becomes invisible to the naked eye. However, even then some light is given off by the few gas molecules that are hit by electrons of sufficient energy, and by long time exposure of a photographic plate the light can be recorded and thus the beam traced. Even in the vacuum that is normally present in electron tubes this light is sufficiently intense to be detected at the same time that space charge conditions are very little disturbed. This is the method that will now be described applied to a practical problem.

THE SPACE CHARGE PROBLEM

The problem at hand was to find the current density and the electron paths in a beam with rectangular cross section being sent out at an arbitrary angle from a plane A (see Fig. 1) at a positive potential towards another plane R parallel to the first and at an arbitrary negative potential. All other methods being inadequate, the photographic method mentioned above seemed to be the only one that could solve the problem. However, although the accuracy of this method supposedly is high it was considered desirable to test a case where the result could be obtained by other means. Consequently the general case was temporarily put aside and a first-order approximation, namely a beam shot out perpendicularly to the planes, was chosen for study. This case was considered to give good enough information for many practical purposes at the same time as the reliability of the method could be checked in two ways. First, with the reflector R at the same positive potential as the emitting plane A, the beam progressed through an approximately field free space, and earlier theoretical results[4–6] could be applied. In case of good correlation

[1] J. T. Wallmark, Proc. Inst. Radio Engrs. **40**, 41 (1952).
[2] A. Wehnelt, Z. physik. chem. Unterricht. **18**, 193 (1905).
[3] E. Brüche and W. Ende, Z. Physik **64**, 186 (1930).
[4] B. J. Thompson and L. B. Headrick, Proc. Inst. Radio Engrs. **28**, 318 (1940).
[5] F. G. Houtermans and K. H. Riewe, Arch. Elektrotech. **35**, 686 (1941).
[6] J. R. Pierce, *Theory and Design of Electron Beams* (D. Van Nostrand Company, Inc., New York, 1949), p. 151.

the experimental results would also serve as a confirmation of these calculations. Second, for the case of the reflector at cathode potential a mathematical solution was worked out and again checked with the photographic method.

EXPERIMENTAL ARRANGEMENT

The experimental electrode system is shown in Fig. 2. The system consists of an essentially flat cathode-grid structure of conventional design and a reflector parallel to the screen grid plane. The width of the electron beam is defined by two aperture plates mounted close to the screen grid. Only the current from one side of the cathode is studied, the current from the other side being collected on a screen electrode.

End effects have been neglected in the following calculations as the length of the cathode is supposed to be large compared to the width of the beam. In the experimental set-up this ratio amounted to a factor of four and consequently end effects were not entirely negligible. These could be studied from the appearance of

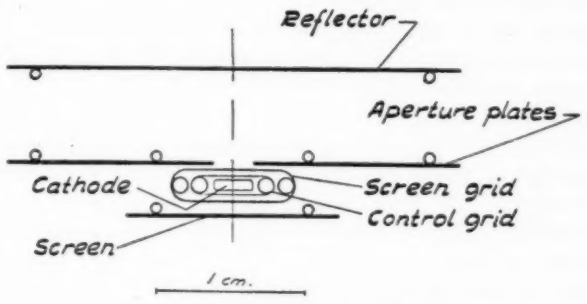


FIG. 2. Experimental electrode system.

burnt spots on the electrodes, probably caused by broken down hydrocarbons from greased seals in the vacuum system. Figure 3 shows such spots on a disassembled system. From the form of such spots it was possible to conclude that end effects do not seriously affect the middle part of the beam, except for returning beams of high current density.

GAS PRESSURE

In all electronic apparatus utilizing an incandescent cathode as electron source the light from the cathode will appreciably mask the light given off by gas molecules. The first step is to minimize this light as far as possible. This can be done by the choice of a rest gas with strong spectral lines in the blue part of the spectrum. Then the use of a blue filter makes it possible to differentiate against the reddish light from the cathode. As it has been shown by Strandell,[7] CO_2 is a suitable gas for this purpose.

The intensity of the light is highest at high pressure but so is the neutralization of the space charge. Thus the best compromise has to be found.

[7] N. Strandell, private communication.

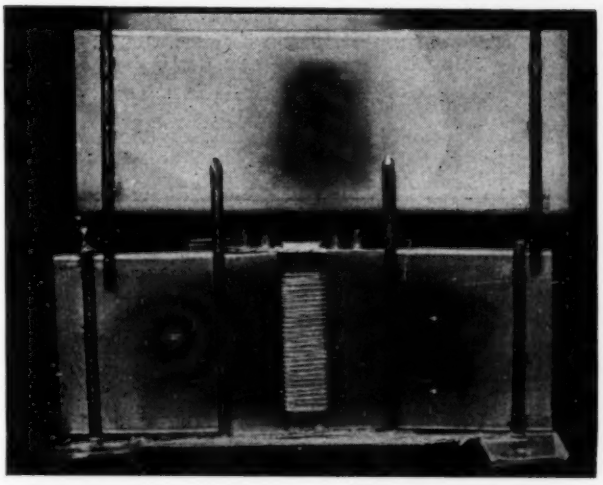


FIG. 3. The electrode system taken apart after some hours use and showing "burnt spots."

Let us investigate then what pressure can be allowed for a certain maximum neutralization, say 10 percent. Let us for simplicity neglect the different distribution in space for positive and negative space charge. Then the ratio of ion space charge to electron space charge is

$$f=\rho_{\text{ion}}/\rho_{\text{el}}=(m_{\text{ion}}/m_{\text{el}})^{\frac{1}{2}}n_{\text{ion}}/n_{\text{el}}, \tag{1}$$

where ρ is the space charge, m is the mass, and n is the number of free ions and electrons, respectively. The probability of ionization for CO_2 in the interval 0–100 volt varies from zero at the ionization potential 14.3 volt to approximately 20 percent at 100 volt. An average figure of 15 percent may be assumed. In terms of electrode dimensions and mean free path, Eq. (1) becomes

$$f=0.15(m_{\text{ion}}/m_{\text{el}})^{\frac{1}{2}}d/l, \tag{2}$$

where d is the traveled distance of the electrons, and l is their mean free path.

The mean free path of an electron is

$$l=(N\pi r^2)^{-1}, \tag{3}$$

where r is the effective radius of the gas molecule ($2.3\cdot10^{-8}$ cm for CO_2), $N=9.68\cdot10^{18}p/T$ is the number of gas molecules per cm^3, p is the gas pressure in mm Hg, and T is the absolute temperature. Then the mean free path for room temperature, say 300°K, is shown by Fig. 4. From these data one can compute the gas pressure for a maximum space charge neutralization of 10 percent to be $p=8\cdot10^{-5}$ mm Hg. Correspondingly a pressure of less than $8\cdot10^{-6}$ mm Hg gives a space charge neutralization of less than 1 percent. At higher pressures, $10^{-3}-10^{-4}$ mm Hg, the space charge of the ions is equal to the space charge of the electrons, corresponding to the phenomenon of thread rays described above. While from a theoretical point of view pressures lower than 10^{-5} mm Hg are desirable to eliminate the influence of ion space charge, the photographic method

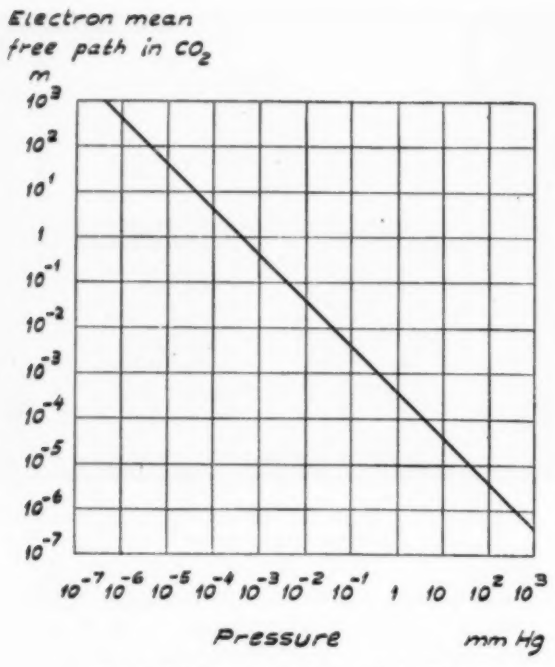


FIG. 4. Electron mean free path in CO_2.

is of course applicable also to higher pressures and to any percentage of positive space charge.

It has to be remembered in the case of electron beams in "field free" systems that the transit time of the ions is considerably increased and consequently the space charge neutralization more severe than for the case treated.

Reducing the pressures below $10^{-5}-10^{-6}$ mm Hg, on the other hand, increases the exposure time to impractical values, as the exposure time varies more than linearly with the vacuum.

Consequently a pressure of $1-7\cdot10^{-5}$ mm Hg was chosen and kept constant with the help of a small controlled leak.

The camera used had a focus distance of 20 cm in order to avoid parallax in the pictures and an opening of $f/2.9$. The exposure time ranged from approximately a minute at high currents to as much as two hours at low currents.

The current was varied by varying the cathode temperature rather than the control grid voltage in order to avoid changes of electrostatic focusing.

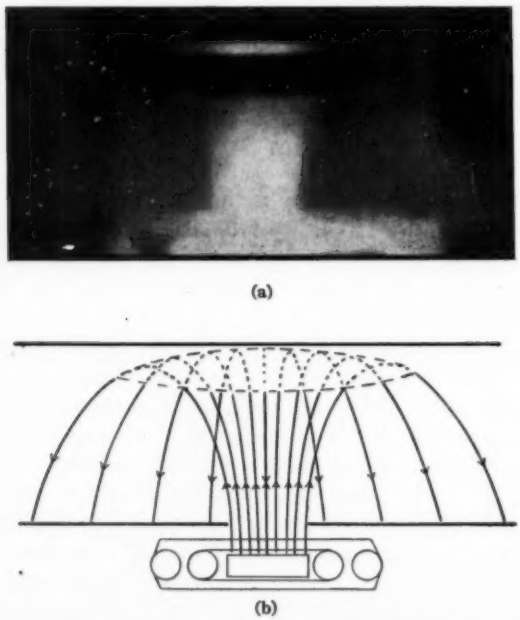


FIG. 5. A typical picture of a reflected beam.

A TYPICAL PICTURE

Let us first discuss what can be seen on a typical photograph such as is shown in Fig. 5. The electrons start with a velocity of 100 volts from the emitting (screen grid) plane. At 100 volts the probability of excitation to higher states is high and rapidly decreases to become zero near the ionization potential (14.3 volts for CO_2). This explains why the light is strong in the beginning of the path, and does not decrease on the returning path although the current density continuously decreases. The excitation probability for the lowest excitation potential (10 volts for CO_2), on the other hand, is very low in the beginning of the path but reaches a sharp maximum at the lowest excitation potential. This shows up as a bright zone nearest to the reflector. For potentials lower than 10 volts, the electron collisions with gas molecules are entirely elastic and no light is given off. In this photograph the current density is high so that the virtual cathode* has retreated considerably towards the emitter, especially in the center of the beam. On the edges the space charge conditions correspond to a lower current density, and the retreat is less pronounced. The bright zone along the aperture plates is caused by light from the cathode that has been reflected or scattered by the end micas.

THE REFLECTOR AT POSITIVE POTENTIAL

A first test of the photographic method was made by keeping the reflector R at the same potential as the emitting plane A. This means that as a first approximation the beam progresses with constant velocity through a field free space.

Consider then a parallel, homogeneous flow of electrons with constant velocity V_A (see Fig. 1), emitted from a plane A perpendicularly toward another plane R on the same potential as and parallel to A on the

* It seems useful to extend the term virtual cathode from the traditional definition: "an emitting plane where the potential and the electric field are both zero" to "an emitting plane where the potential is zero." The first definition is analogous to a thermionic cathode with space charge limited current, the second definition is analogous to any thermionic cathode, even with temperature limited current. This has been suggested earlier by Salzberg and Haeff (reference 9) and others.

distance d. Assume a cross section of the beam with dimensions $2Y_A \times l (l \gg Y_A)$.

With the coordinate system of Fig. 1 the equations of electron motion are

$$\begin{cases} \dfrac{d^2V}{dy^2} = -\rho/\epsilon_0 & \text{Poisson's equation} \quad (4) \\ i = -\rho v & \text{continuity equation} \\ mv^2/2 = eV & \text{energy equation} \\ \dfrac{d^2Y}{dt^2} = -eE/m & \text{force equation.} \quad (5) \end{cases}$$

From (4) the electric field strength at the edge of the beam can be integrated:

$$E = -i_A Y_A/\epsilon_0 (2eV_A/m)^{\frac{1}{2}}.$$

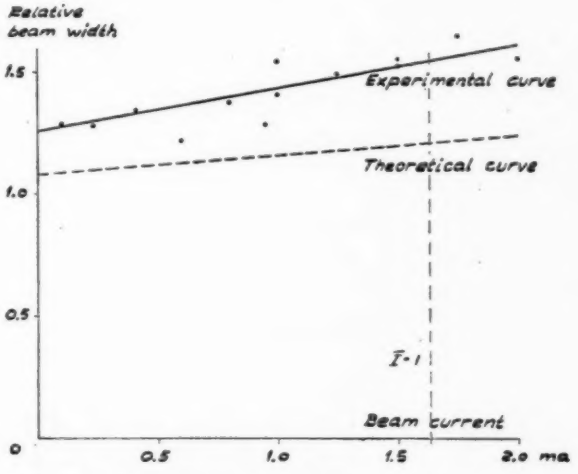


FIG. 6. Spread *vs* current of a constant velocity electron beam.

Combination with (5) gives the solution

$$Y = Y_A[1 + ex^2 i_A Y_A/2m\epsilon_0 (2eV_A/m)^{\frac{1}{2}}].$$

As usual in space charge calculations introduction of the relative quantity:

$$\bar{I} = i_A/4\epsilon_0 (2e/m)^{\frac{1}{2}} V_A^{\frac{1}{2}} (9x^2)^{-1},$$

gives a much simpler expression:

$$Y = Y_A[1 + \bar{I}/9].$$

$\bar{I}$ is the relative current compared to space charge saturated current between planes at a distance x and with the potential difference V_A. This equation represents a straight line and has been drawn in Fig. 6 with practical data inserted and with the effect of initial velocities included.

Some experimental results are shown by the series of photographs in Fig. 7. The beam width at the reflector is plotted in Fig. 6.

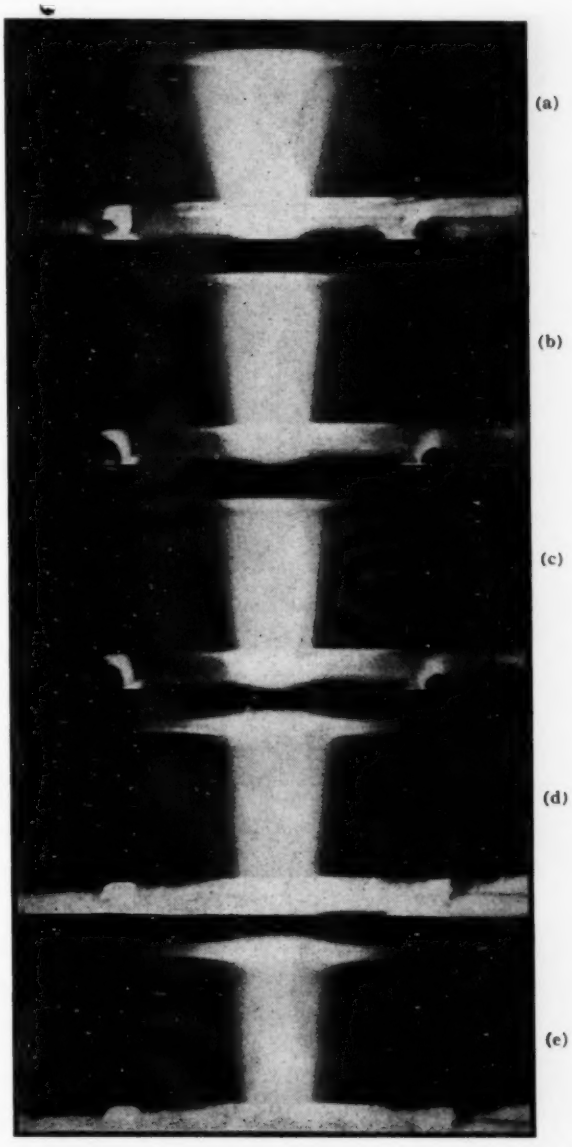
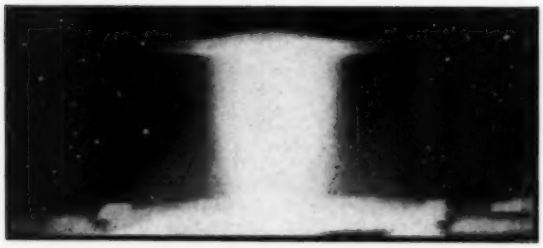


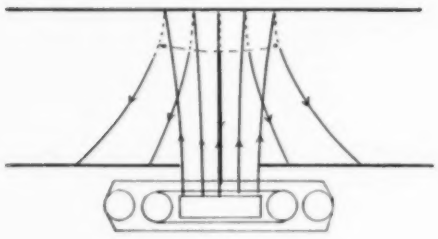
FIG. 7. Photographs of a constant velocity beam. Beam potential 100 v. Pressure $7 \cdot 10^{-5}$ mm Hg. Beam current (a) 2 ma, (b) 1.5 ma, (c) 1.0 ma, (d) 0.5 ma, (e) 0.1 ma.

The experimental and theoretical curves have the same form but two differences occur. First the experimental curve has a higher slope than the theoretica- curve. This is probably caused by the space charge del pression of potential in the x direction and the consequent increase in transit time with current, which had to be neglected in the calculations. Second, the experimental curve shows a spread approximately three times the spread caused by initial velocities. This may be contributed by the aperture plates which act as a spreading lens.

A rather special case is shown by Fig. 8 where the reflector is at a moderately positive potential and consequently gives off secondary electrons which are collected by the aperture plane. The secondary electron stream of course increases beam spread.



(a)



(b)

FIG. 8. Photograph of a retarded beam releasing secondary electrons. Pressure $7 \cdot 10^{-5}$ mm Hg. Beam current 1.8 ma.

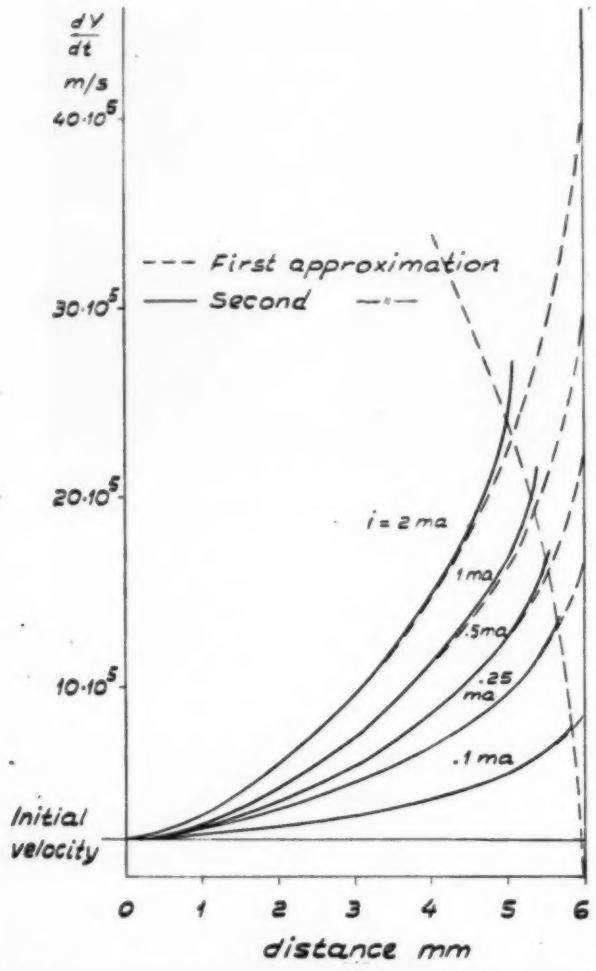


FIG. 9. Electron velocity dV/dt *vs* distance for determination of the turning point. Dotted lines show first approximation, solid lines second approximation.

THE REFLECTOR AT CATHODE POTENTIAL

This case has been treated earlier by Moss,[8] who, however, considered circular beams. Consider a parallel, homogeneous flow of electrons with a velocity V_A, emitted from a plane A at the potential V_A, perpendicularly toward another plane R, parallel to A at the distance d, and of zero potential. Assume a cross section of the beam with dimensions $2Y_A \times l (l \gg Y_A)$. The

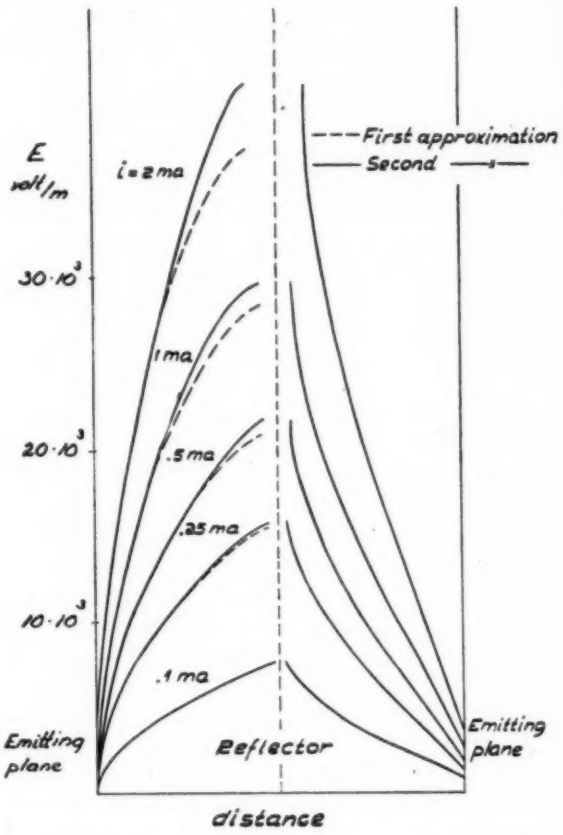


FIG. 10. Electric field strength at the edge of the beam *vs* distance.

equations of motion are

$$\begin{cases} \dfrac{\partial^2 V}{\partial x^2} + \dfrac{\partial^2 V}{\partial y^2} = -\rho/\epsilon_0 & \text{Poisson's equation} \\ i = -\rho v & \text{continuity equation} \\ mv^2/2 = eV & \text{energy equation} \\ \dfrac{d^2 Y}{dt^2} = -e/m \cdot E & \text{force equation.} \end{cases} \tag{6}$$

Let us further assume that V is linear in x. Then elimination of v and ρ and integrations give

$$E = -[4 i_A Y_A (V_A^{\frac{1}{2}} - V^{\frac{1}{2}})/\epsilon_0 Y (2e/m)^{\frac{1}{2}}]^{\frac{1}{2}}.$$

[8] H. Moss, Wireless Engr. 22, 316 (1945).

Combination with (6) gives

$$\frac{d^2Y}{dt^2}=(2e/m)^{\frac{1}{2}}[i_AY_A(V_A^{\frac{1}{2}}-V^{\frac{1}{2}})/\epsilon_0Y]^{\frac{1}{2}}. \qquad (7)$$

Equation (7) has been solved by an iterative method, whereby first V and Y have been assumed constant, and the consequent change in Y after a small distance has been calculated from (7). This new value of Y together with a new value of V have been used for the next distance, etc. The turning point has been obtained from a plot of dY/dt *versus* distance as the point where the velocity in the y direction is equal to the potential at that point. This is shown in Fig. 9. Figure 10 shows the

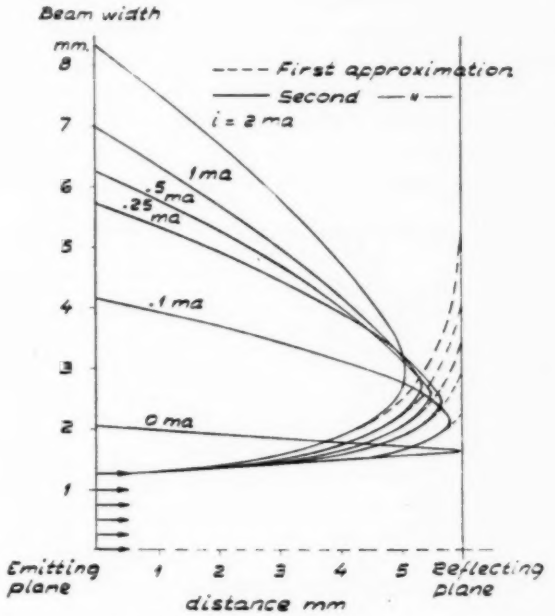


FIG. 11. Electron paths. Dotted lines show first approximation, solid lines second approximation.

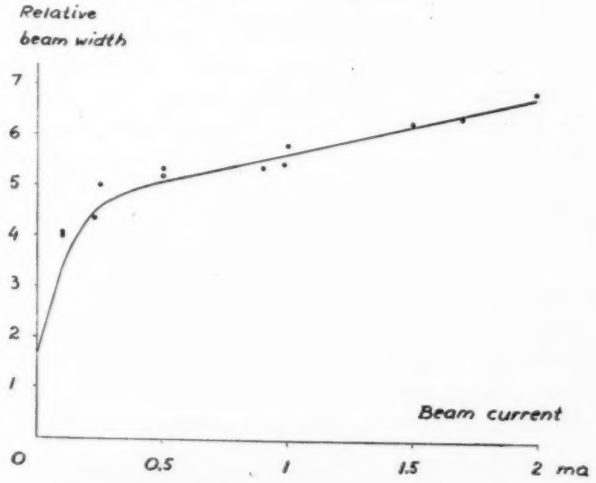


FIG. 12. Spread of reflected electron beam. Points represent experimental values, curve theoretical values.

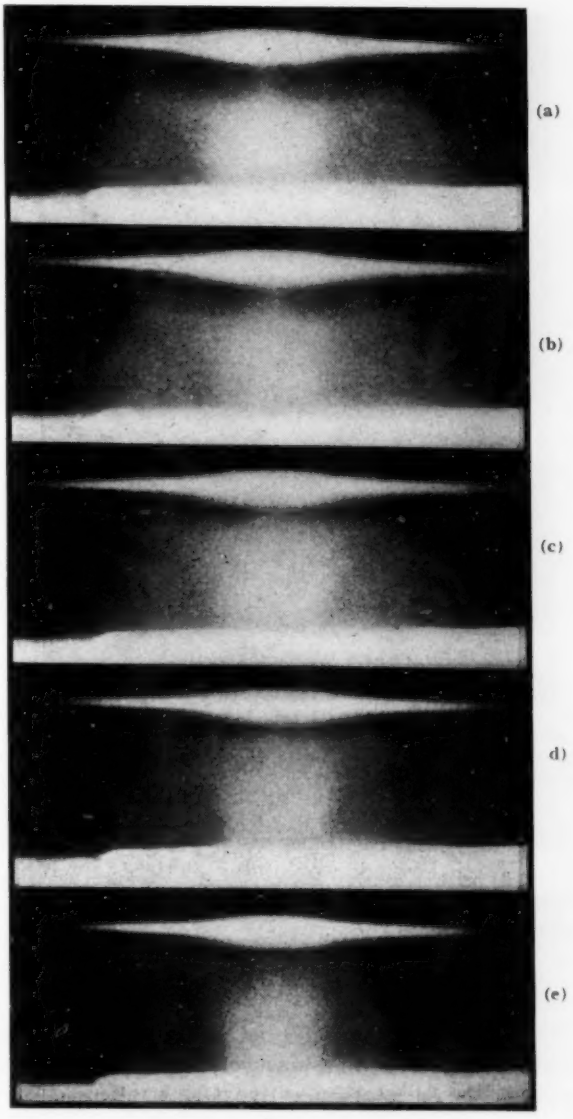


FIG. 13. Photographs of a reflected beam. Pressure $1\cdot10^{-5}$ mm Hg. Beam current (a) 2 ma, (b) 1.5 ma, (c) 1.0 ma, (d) 0.5 ma, (e) 0.1 ma.

electric field strength at the edge of the beam *versus* distance for different current densities. After a first calculation and determination of the turning point, a second approximation is made by recalculating the spread, taking into consideration that $v=(v_x^2+v_y^2)^{\frac{1}{2}}$. The resulting electron paths are shown in Fig. 11 again with current density as parameter. The resulting spread is plotted in Fig. 12 as a solid curve.

Some experimental results are shown to the series of photographs in Fig. 13, and from measurements on such photographs the points in Fig. 12 have been plotted.

The theoretical curve fits the experimental points very well except that the spread is again somewhat larger than calculated. This would be even more so especially at low currents if the current values were in-

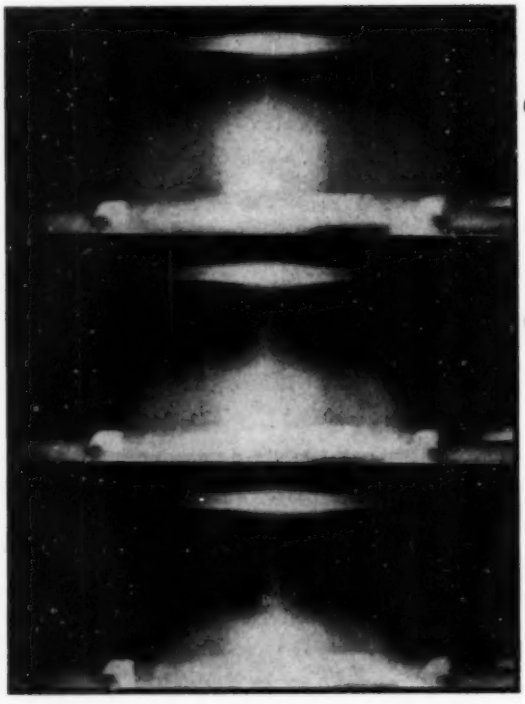


FIG. 14. Photographs of a reflected beam with reflector voltage negative (a) -10 v, (b) -50 v, (c) -100 v. Pressure $7 \cdot 10^{-5}$ mm Hg. Beam current 1 ma.

creased to take into consideration also the returning beam. Then the slope would also be higher for the experimental curve than for the theoretical curve for the same reason as in the foregoing case.

THE REFLECTOR AT NEGATIVE POTENTIAL

The electrode system did not allow longer beams to be tested. Shorter beams, however, could easily be made by giving the reflector a negative potential. Some results are shown in Fig. 14. The peak towards the reflector is formed by ions. The gradual retreat of the virtual cathode closely follows theory.[9–11]

Figure 15 shows measured values of the widening of the returning beam *versus* beam length. The figures have been normalized by division with initial beam width.

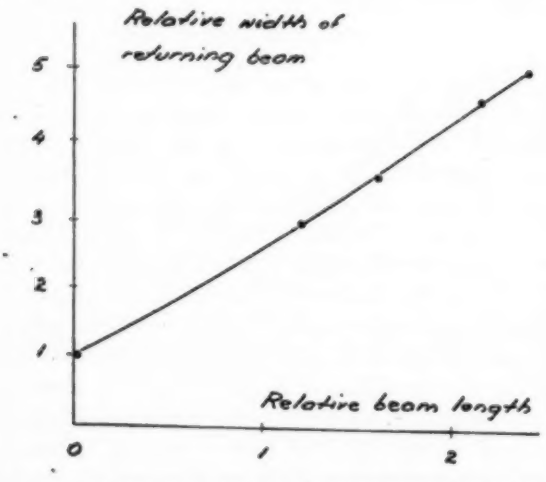


FIG. 15. Relative width of returning beam *vs* relative beam length.

RETROGRADE MOTION OF THE VIRTUAL CATHODE

As can be seen from the photographs in Fig. 13 the beam seems to get shorter for higher currents, not only on the edges where the retrograde motion is due to a gradual pick-up of dynamical energy in the sidewise direction, but also in the center of the beam. This is due to a retrograde motion of the virtual cathode at the reflector—the turn around point of the electrons. This retrograde motion has been calculated for the simple case of an infinitely wide beam.[9–11] Assume a contact potential difference between the cathode and the reflector of 1 volt. Then the position of the virtual cathode as a function of current can be calculated. Knowing this position the equipotential line $V=10$ volts can also be calculated and is shown in Fig. 16 as a broken line. The experimental curve does not fall as rapidly as the theoretical curve for two reasons. First, the spread of the beam causes the current density to change much slower than the beam current. Second, the returning beam adds to the current density although the magnitude of this effect is difficult to determine. An approximate calculation brings the theoretical and experimental curves very close to each other.

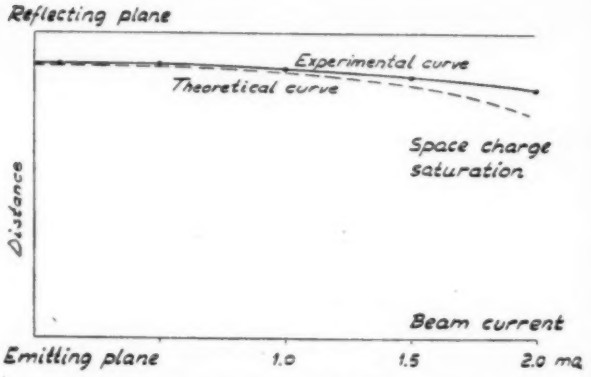


FIG. 16. Retrograde motion of 10 v equipotential line *vs* current. Dotted line theoretical, solid line experimental values.

CONCLUSION

The data given here on the spread of reflected electron beams are valid for a certain width-to-length ratio of the beams. The method may be applied with good accuracy also to narrow beams, i.e., beams with a lower width-to-length ratio, where the assumption of V as a

[9] B. Salzberg and A. V. Haeff, RCA Rev. **2**, 336 (1938).
[10] Fay, Samuel, and Shockley, Bell System Tech. J. **17**, 49 (1938).
[11] J. T. Wallmark (to be published).

linear function of x would hold. However, in that case the retreat of the virtual cathode, which is a function of current density, would appear first at higher beam currents.

For broad beams, on the other hand, the retreat of the virtual cathode can be easily calculated; but there additional spread of the beam would appear near space charge limited current density, due to increased transit times of the electrons. Or in other words V as a linear function of x would be a poor approximation near space charge limited current density. The accuracy of the iterative calculation method is of course dependent upon the number of successive steps involved. In this case twelve steps were used, and recalculating a third approximation on the most unfavorable case (high beam current) showed less than 10 percent deviation in beam width. As this is within the accuracy of the assumptions made, further calculations do not improve the over-all accuracy. From the very good agreement between theoretical and practical results it is concluded that the photographic method is a very useful tool in complicated space-charge problems of static or quasi-static nature.

Even the interior of beams or space charge systems can be investigated as long as a contrast can be created, e.g., by preventing the emission over a small area of the cathode thereby introducing a dark part in the beam. It is of course also possible to investigate space charge clouds in which the electron velocity is below 10 volts if a small search beam having a velocity greater than 10 volts is used and photographed. Special care has to be taken in places where trapping of ions may take place.

The accuracy of the photographic method is mainly limited by photographic difficulties such as contrast, parallax, etc., and the difficulty of opening up the interior of vacuum tube systems to the line of sight without disturbing the electrical conditions. The intrinsic accuracy of the method itself can be expected to be better than any other method known.

ACKNOWLEDGMENT

Many of the practical details for the work were pioneered by N. Strandell, and for most of the photographs the assistance of S. Hellström is gratefully acknowledged.

JOURNAL OF APPLIED PHYSICS VOLUME 24, NUMBER 5 MAY, 1953

Metal Capillary Cathodes

H. Katz
Siemens and Halske Tube Laboratory, Munich, Germany
(Received November 17, 1952)

Since 1941 we have been working on the development of new types of cathodes, using elements such as barium or thorium as emitters. If provided with a suitable base, these elements are able to migrate and, acting as an atomic layer, reduce the work function of the base material. By introducing a separate supply source below the emitting surface, a work-function-reducing substance will be caused to migrate to the cathode surface through small capillaries. Evidently a wide range of new types of cathodes is feasible, since the independently stored emitter of a metal capillary cathode can be held at a temperature different from that of the cathode surface. The subject of this paper is the description of some of these cathodes.

I. INTRODUCTION

THE principle of metal capillary cathodes may be understood from the following example.[1] The crosses in Fig. 1 represent a source of material which will yield barium at high temperatures. This source is confined to a metal cup "M" which contains a porous disk "T" tightly welded to the rim. Therefore, as the source evaporates, the barium can escape only by passing through the pores of the disk. Both cup and disk are heated by a heater "H." By means of evaporation, free barium from the source is deposited on the back of the disk. From there it reaches the surface through capillaries, where it forms a film which acts to reduce the work function of the tungsten. If the rate of evaporation of barium from the source is kept within proper limits, and if at the same time it is high enough to replace the amount lost from the surface, a steady rated emission can be expected over a long period. This period may be extended by providing a correspondingly larger source.

A metal capillary cathode will operate properly under the following conditions:

(1) The back of the cover of the cathode must be within reach of a substance reducing the work function continuously spent by an independent source.

(2) Capillaries have to be formed through which the active material can get through to the cathode surface.

(3) The cathode has to stay covered by the material that is coming through the capillaries.

(4) The rate of evaporation from storage has to correspond to the rate of loss from the activated surface.

[1] See French Patent Specification Nr. 903976, published October 23, 1945; German priority August 19, 1942.

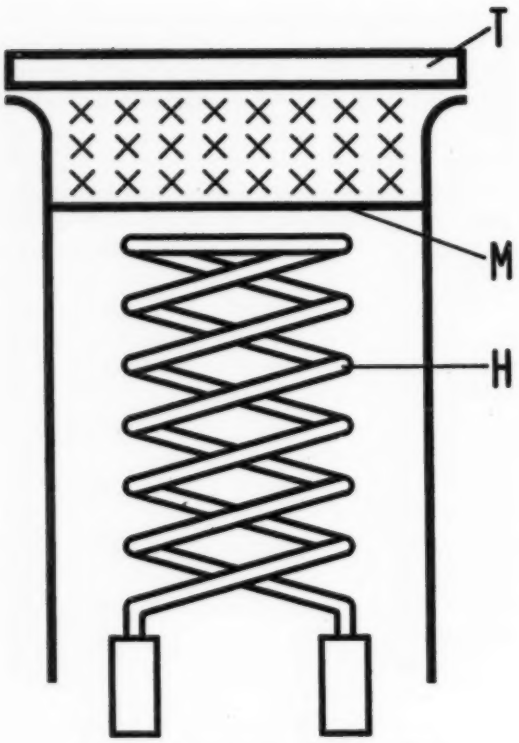


FIG. 1. Basic structure of a MK cathode.

These conditions can be fulfilled in a variety of ways. We shall examine the individual points more closely.

(1) Elements which reduce the work function of tungsten are such as barium, thorium, cesium, lanthanum, cerium, yttrium. Thorium, for example, may be evaporated from a source of thoria or, otherwise, directly from metallic thorium.

(2) Capillaries may be formed in various ways. Either the cover of the cathode should be made of wires or bands abutting on one another, or as an alternative, of a porous material.[2]

(3) The coating of the cathode surface can be effected by migration. A suitable material and a suitable base must be provided for this purpose. Thorium is known to migrate easily on tungsten and molybdenum.

(4) The rate of evaporation from the source can be regulated by choice of materials used; for instance, at the same given temperature less thorium will evaporate from thoria than from metallic thorium. This is due to the fact that thoria has to be reduced first. Given a source of the same material, however, the rate of evaporation can be considerably modified by regulating the temperature, since, if metal capillary cathodes are used, there is no need to keep both source and cathode surface at the same temperature.

It is felt necessary to give the new cathode type a common designation with an addition to distinguish them from each other. MK is to characterize its fundamental design and stands for "metal capillary."[3] The remainder of the designation refers to the base material and to the source material filled into the cup. The formula MK-W-Ba(Sr)CO_3 thus signifies a metal capillary cathode with tungsten as base material and Ba(Sr)CO_3 as stored emitter. The cathode attracting top interest at present will be the type containing a porous base material. As an example we will therefore proceed to discuss the production and characteristics of porous tungsten.

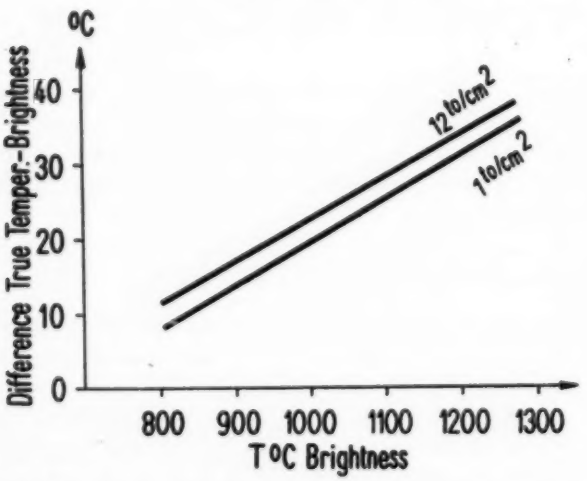


FIG. 2. Brightness *vs* true temperature—brightness ($\lambda = 0.65\mu$) of porous tungsten. True temperature was measured by producing a number of holes (diameter 0.3 mm, depth 2 mm) in compressed disks of 3-mm thickness. Increased pressure generally yielded increased difference between true temperature and brightness.

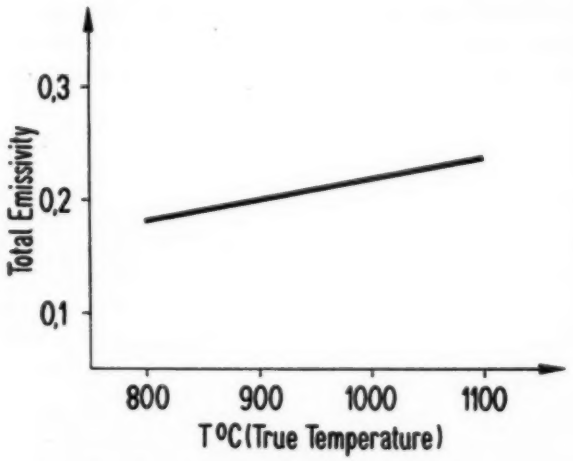


FIG. 3. Total emissivity of porous tungsten (approximate measurements).

II. POROUS TUNGSTEN AND SOME OF ITS CHARACTERISTICS

Porous tungsten is, of course, easier to produce then a dense material. A heavy suspension of powder in butylacetate is poured into a mold and compressed as

[2] H. Katz, "Metal capillary cathodes, a new type of cathode." Siemens and Halske AG., Research Bulletin, II, 123–126 (1951).
[3] Capillary in German is Kapillar.

the purpose may require with pressures up to 16 tons per cm². Sintering of the molded body is thereupon effected in purified hydrogen at temperatures close to 2000°C.

Figure 2 shows our measurements of true temperature *versus* brightness of our porous tungsten. The pressure is of little influence. The values of total emissivity as shown in Fig. 3 are derived from the heating power which had to be applied to cathodes enclosing the heater as far as possible, taking into account an estimated value for the heat by-pass. It may be noted that for the same temperature, values for porous tungsten do not exceed those for BaO+SrO layers of customary thickness on nickel.

III. EXAMPLES OF MK CATHODES

A. The Tungsten-Barium-Strontium-Carbonate: MK-W-Ba(Sr)CO_3 Cathode

This cathode was the first type we arrived at in applying some of the principles set forth in our introduction. Since the properties of this cathode—under the designation "*L*" cathode—are being frequently discussed beginning with a comprehensive paper;[4] we think it unnecessary to go into close details. Considering that the cathodes described below are more interesting today than this MK-W-Ba(Sr)CO_3 cathode, it may be sufficient to include a sketch of this original model (Fig. 4). It consists of a cup made of thin pressed molybdenum sheet arc-welded to the porous tungsten disk. A notch stamped into the rim of the Mo cup is left open temporarily for pouring in the Ba(Sr)CO_3, and this rim is closed later on by welding. A particularly durable model is shown in Fig. 5.

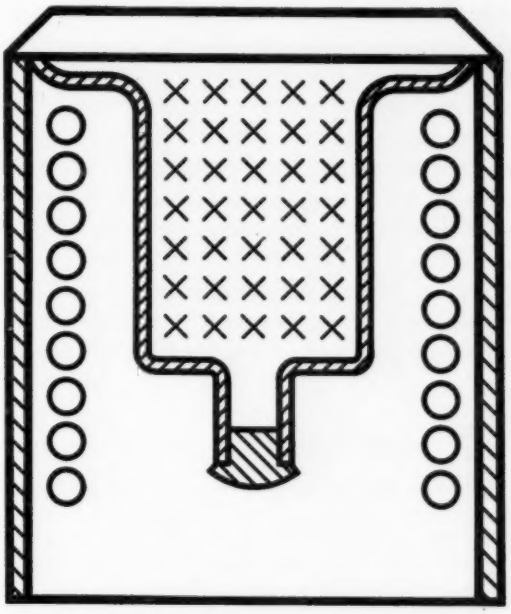


FIG. 5. MK-W-Ba(Sr)CO_3 cathode for large storage of emitting material.

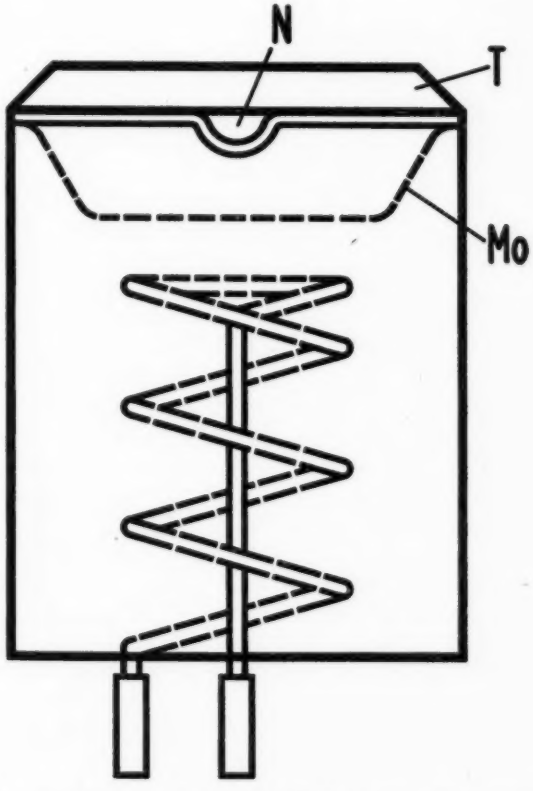


FIG. 4. Structure of MK-W-Ba(Sr)CO_3 cathode. T=porous tungsten disk, Mo=molybdenum container, N=gap, initially left open for pouring in the stored material. The container Mo is arc-welded on the tungsten disk in such a way that the pores at the side of the tungsten disk are closed.

[4] Lemmens, Jansen, and Lossjes, Phillips Tech. Rev. **11**, 349 (1950).

In our conception of the operation of an MK cathode it is immaterial from which source the substance is supplied which serves to reduce the work function of the cathode surface. Barium-strontium-carbonate may be only one among a number of possible materials. In future, a compound will be sought which yields barium without previously freeing large quantities of gas. The number of compounds feasible for such use will be substantially increased when recourse is taken to the possibility of holding source and cathode surface at separate temperatures.

B. The Tungsten-Thoria: MK-W-ThO_2 Cathode

The design principle of this cathode is the same as of the MK-W-Ba(Sr)CO_3 cathode, except that barium-strontium-carbonate is replaced by thoria as source of emitting material. Because of the high operating temperature the molybdenum also has to be replaced by another material. However, as this cathode is only intended for experimental use, it may be constructed in the manner that thoria is sprayed on a tungsten wire; and this combination is closely surrounded by a coil of tungsten wire (0.03 mm). This cathode is heated

directly. Its behavior has previously been described in detail.[5] After the formation process, at a temperature around 2000°C, it acquires a high specific emission. Of far greater interest, however, are the inertia effects occurring during a change of temperature. This may lead to the conclusion that good emission might also be obtainable at lower temperatures if more thorium could be freed from the thoria source to cover the cathode surface. It is true that the amount of thorium can be increased by raising the temperature of the cathode, but in this case the evaporation rate of the thorium at the surface would also go up. As a result, the gain in

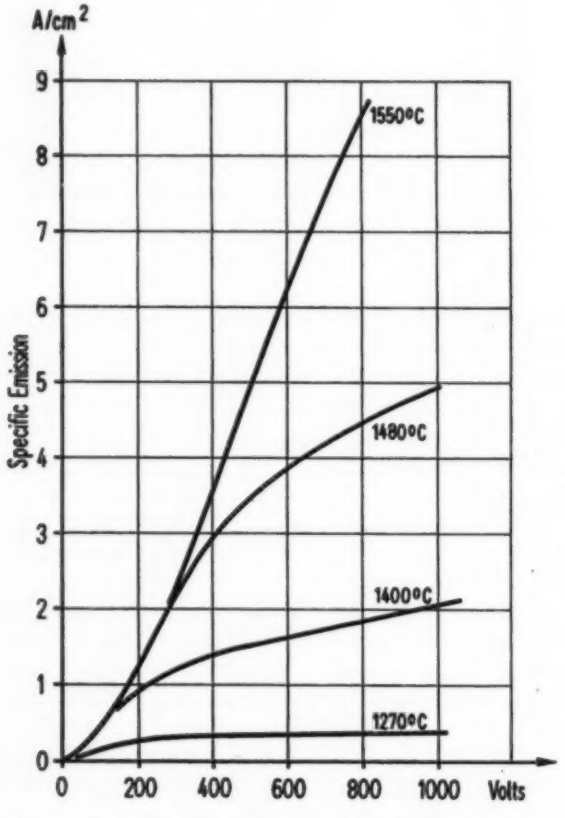


FIG. 6. Specific emission (pulsed) of a MK-W-TH cathode. Parameter true temperature in centigrade.

emission would be achieved only at the expense of life and an intolerably high amount of heating power. To attain our ends we may once again attempt to keep the source at a higher temperature than the surface. This can certainly be done by structural modification. The method we have adopted, however, has been to increase the rate of reduction, or, more precisely, to avoid reduction by employing metallic thorium as a source. This brings us to the cathode to be described next.

[5] H. Katz and K. L. Rau, Frequenz 5, 192 (1951).

C. The Tungsten-Thorium: MK-W-Th Cathode

The migration of thorium on tungsten was first recorded in a paper by Brattain and Becker.[6] Other authors[7] subsequently arrived at the same conclusion after electron optical means had been applied for a direct observation of this migration.[8] Migration is, therefore, no longer in doubt. The paper by Brattain and Becker also provided insight on the required temperature. Migration was observed to start at 1535°K and is considerably accelerated at 1655°K. The proof of migration makes it unnecessary to search for another explanation for the functioning of cathodes, and it permits a reliable prediction of the behavior of a cathode under various operating conditions. The data on temperature governing the migration of thorium on a base material also refer to the temperature to be used for the formation process, and to the minimum working temperature.

We have based our first experiments with MK-W-Th cathodes on a construction similar to that shown in Fig. 4. The cup contains metallic thorium. The tungsten disks have been made according to a procedure described in a former publication[2] by sintering tungsten powder with molybdenum powder and extracting molybdenum by acid. Heating is effected by electron bombardment from an auxiliary cathode. Activation of the cathode is performed by operation at 1700°K after normal outgassing procedures of tube elements. Once the anode voltage is applied, emission continually increases and soon attains its final value. The formation process has been continued for several hours after removing the tube from the pump.

The cathode attains good emission at temperatures appreciably lower than those required by the MK-W-ThO_2, or by the ordinary thoriated tungsten cathode. Figure 6 shows emission values of an average sample at various temperatures. Measurements of the work function φ and the constant A, according to Richardson's equation $J_s = AT^2 e^{-11\,605/T}$, have proved that φ nearly agrees with the values already known for thorium on molybdenum and tungsten, the constant A, however, reaches much higher values. At the best model we have had so far[4] 100 $A\ cm^{-2} \cdot deg^{-2}$. The emission obtained with this MK-W-Th cathode is in accordance with the predictions based on the behavior of the MK-W-ThO_2, to which reference has already been made.

Further cathodes of this kind, with indirect heating and tungsten disks produced without molybdenum, are being tested. Difficulties may arise by the diffusing of thorium through the dense cup material at high temperatures. Of further interest is the fact that the carburization customary in the case of thoriated tungsten cathodes has not been performed with this cathode.

[6] W. H. Brattain and J. A. Becker, Phys. Rev. 43, 428 (1933).
[7] M. Benjamin and R. O. Jenkins, Phil. Mag. 26, 1049 (1938).
[8] E. Brüche and H. Mahl, Z. Tech. Physik. 16, 623 (1935).

It may be possible that carburization is advantageous with a required high specific emission and a corresponding increase of cathode temperature.

D. The Carbon-Thorium: MK-C-Th Cathode

Although this cathode, with porous carbon[9] replacing porous tungsten, does not greatly differ from others described here, its existence is perhaps of some interest. Structurally it is rather similar to the MK-W-Th described under C. Activation is more complex and requires more time, but the ultimately obtained emission is also good. This cathode also possesses high operation stability.

E. The Tungsten-Barium Beryllium and the Molybdenum-Barium Beryllium: MK-W-(BaBe) and MK-Mo-(BaBe) Cathodes

The principle of holding source and cathode surface at different temperatures will now be elucidated for example by the cathode mentioned above. In the case of replacing thoria by thorium, we have been guided by the desire to avoid reduction and provide a larger amount of thorium for coating the cathode surface. We now follow the same principle in replacing barium-strontium-carbonate or a similar compound by barium. With the MK-W-Th cathode it was found that as a rule no special means are required to adjust the rate of evaporation of the metallic thorium to the rate of consumption at the surface. This does not apply to barium. If the operating temperature of a cathode is assumed to be, for example, 800°C, the metallic barium would, evaporate too quickly for the cathode to work. Metallic barium will, of course, not be used in any case because of its instability. But there are alloys such as BaAl or BaBe in which the barium is fairly stable. The latter above all possesses the additional advantage that its barium content evaporates at a higher temperature than pure barium. Its rate of evaporation would, however, still be too high, and, therefore, the source must be spatially separated from the cathode surface, and the heater placed in a manner that the source could remain at a lower temperature. A model of this cathode is shown in Fig. 7.

The amount of barium evaporating to the cathode surface will depend on how large a source is used and how high a temperature is applied. Long life is secured if the storage is large and its temperature low. In the course of time, and even when the temperature of the source remains constant, the yield of barium will diminish, but if a tube has been sufficiently aged, disturbing influences and consequently the requirement of barium will also diminish.

The adjustment of the source temperature with respect to that of the cathode surface is, of course, a matter of spacing. The temperature of the source can

[9] Porous carbon disks of high purity have been produced for us by the Schunk and Ebe Company at Giessen.

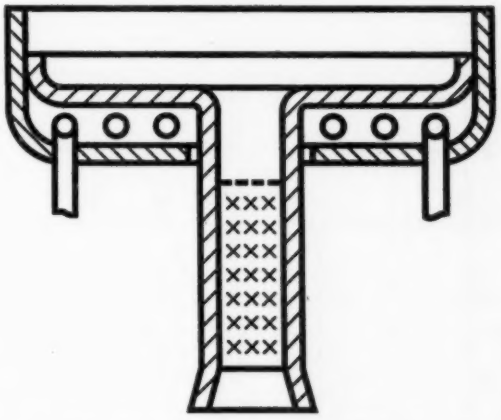


FIG. 7. Basic structure of a MK-W-BaBe cathode.

be regulated by conductivity, thickness, length, and radiation characteristics of which its container is made and which connects it to the cathode.

The cup should preferably be held at a somewhat higher temperature so that an adequate yield of barium is assured. If highly porous tungsten disks are used for cathodes, a barium deposit may possibly be formed within the tube, but its growth will soon decrease. In this connection tungsten disks of greater density have proved highly suitable. These, manufactured at a pres-

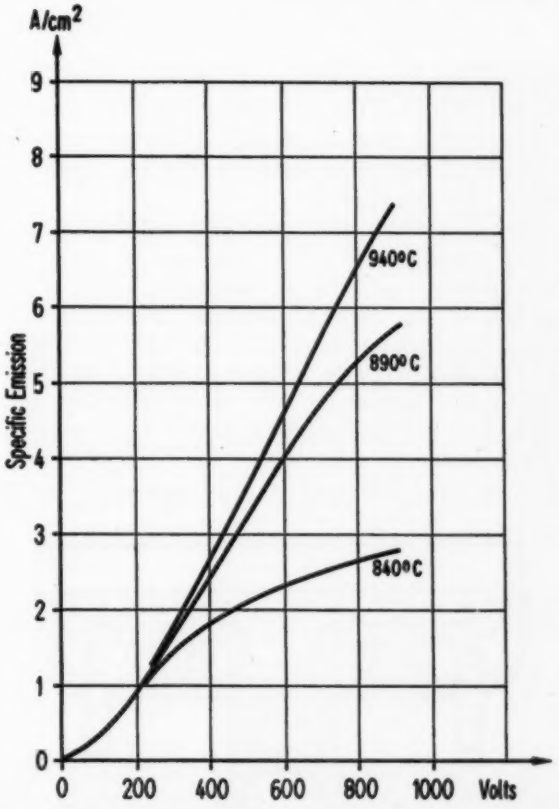


FIG. 8. Specific emission (pulsed) of a MK-W-BaBe cathode of good quality. Parameter true temperature in centigrade.

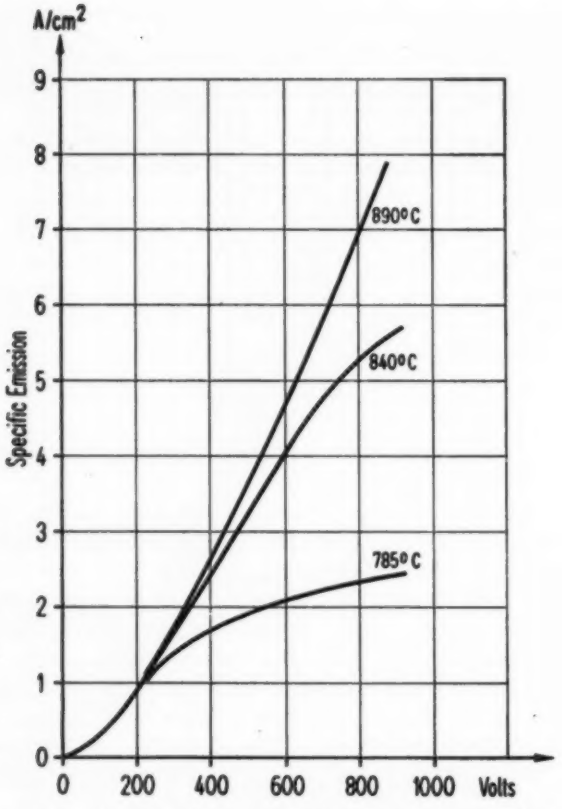


FIG. 9. Specific emission (pulsed) of a MK-Mo-BaBe cathode. Parameter true temperature in centigrade.

sure of 16 tons per cm² at high sintering temperature, will diminish the deposition of barium on tube components even if the temperature of the source is somewhat high. A slightly prolonged activation period would not be a disadvantage.

The construction of this cathode assures the supply of any required amount of barium to the surface. Consequently, as seen in comparing MK-W-ThO_2 and MK-W-Th, the MK-W-(BaBe) also attains more interesting emission values at lower temperatures than the MK-W-Ba(Sr)CO_3. Figure 8 shows the pulsed emission of a cathode of this type. The lower working temperature and the resulting lower heating power are, however, not the only merits of this cathode. The activation, for instance, is also appreciably simpler due to the nonoccurrence of chemical reactions forming gases which would have to be pumped out. Depending on the porosity of the tungsten disk, full emission is obtained within a period between 10 minutes and 1 hour at an activation temperature of 1100°C. After sealing off the tube, the cathode once more will have to be aged by several hours.

Besides constructive simplification partly based on the usability of nickel there is above all the possibility of using porous molybdenum as disk material. This has favorable emitting properties (see Fig. 9). Our life tests are directed to ascertain whether the life of this cathode may not be limited by the molybdenum losing its porosity also at low temperatures.

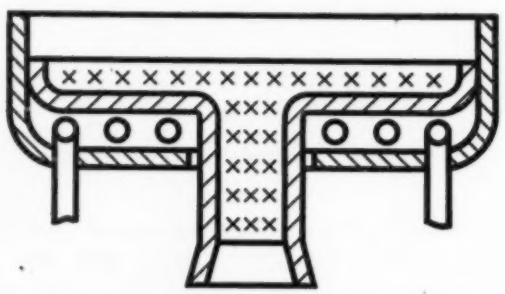


FIG. 10. Basic structure of a MK-W-$BaCO_3$+Si cathode.

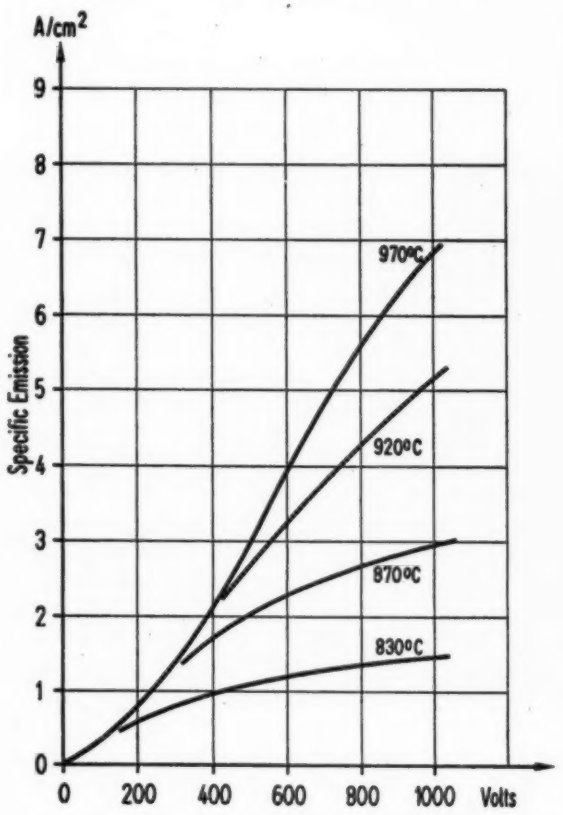


FIG. 11. Specific emission (pulsed) of a MK-W-$BaCO_3$+Si cathode. Parameter true temperature in centigrade.

F. The Tungsten-Barium-Carbonate-Silicon: MK-W-$BaCO_3$+Si Cathode

The MK cathode with barium-beryllium as Ba-source will not be advantageous in all cases despite its good properties. This may be the case if there are special, necessary constructive means in tubes, which make it difficult to maintain the different temperature of the cavity and the surface of the cathode. Therefore, we have devised a MK cathode (Fig. 10), which has the same temperature everywhere, but still shows good emission properties of the MK-W-(BaBe) cathode.

We used the knowledge that it is of utmost importance to supply the cathode with a sufficient amount of barium. This necessary amount cannot be obtained using barium-carbonate only. But there are admixtures accelerating the reduction of the barium-carbonate. Many elements are usable therefore, such as carbon, titanium, magnesium, and others. We used silicon, adding it to the barium-carbonate in coarse-grained form. The now increased production of free barium decreases the cathode's operating temperature. Figure 11 shows that with this cathode one can get about as good an emission as with the MK-W-BaBe cathode. An operating temperature between 850 and 900°C ought to be sufficient in most cases. It may be advantageous to replace $BaCO_3$ by BaO because the reaction products of the freed gases with tungsten sometimes diminish the emission.

FIG. 12. Electron image of a MK cathode with pencil inscription.

The emission curves for all MK cathodes in this article refer to such temperature ranges, promising a long life of the cathode in vacuum tubes. The life of these tubes will primarily be determined by the barium evaporation. This evaporation will be essentially lower if the cathode works in gas filled tubes. Such MK-W-$BaCO_3$+Si cathodes have been working already for quite a while in our rectifier tubes for emission tests, at temperatures of about 1100°C. Peak currents up to 50 A/cm² are drawn in 50 cycle operation, and even at 150 A/cm² there are no signs of saturation of the cathode.

IV. CONCLUSION

We believe to have shown with the examples of MK cathodes discussed in this article that a sufficient amount of substance from the storage has to be provided for the coverage of the surface in order to achieve large emission at low temperatures. But on the other hand, the subsequent delivery should not essentially exceed the amount necessary for optimal coverage. The disappearance of the coating on the cathode surface can be diminished by reducing evaporation from it, therefore, a low operating temperature is advantageous. Furthermore, the coverage of the cathode is a function of the migration speed on the surface. Consequently the best cathode is the one showing lowest evaporation and highest migration speed of the work function reducing substance on the surface of the cathode. Whether or not tungsten is already an ideal base for migration of barium must be left undecided.

TABLE I.

Type of cathode	Working temperature true temperature in °C	Specific emission (pulsed) in A/cm²
MK-W-Th	1500	3
MK-W-Ba(Sr)CO_3	1000	3
MK-W-BaBe	900	5
MK-Mo-BaBe	850	5
MK-W-$BaCO_3$+Si	900	4

Other substances, which may be more favorable in terms of evaporation and migration, will not always maintain their porosity at the required temperatures. However, one may try porous tungsten as substratum material for other layers with perhaps better properties. A simple experiment we have made in this respect was to write with an ordinary pencil "MK-Kath" on the porous tungsten disk, and to produce the electron image of this cathode (Fig. 12). It is very surprising to see bright areas forming themselves around this pencil line, which represent a considerable increase of emission on the hardly visible cathode surface. If the cathode already is entirely dark in consequence of lowering its temperature, these areas are still considerably bright.

While these experiments showed a distinct increase of emission at certain areas on the cathode, other experiments have not shown a clear result. Like Ryde and Harris[10] we have "poisoned" fully seasoned cathodes by admitting air. Thereafter, these cathodes were rebaked. The second process of formation takes somewhat longer than the first one, but the result is a stable cathode, the emission of which is hardly different to that of non-oxidized cathodes. Reimann[10] attributed to the W-O-Ba system the emission from a tight layer of barium atoms on a tungsten filament, which was the result of the same treatment. Additional oxidation did not affect a distinct increase of emission in our experiments.

See Table I which shows specific emissions that can be expected from the various MK cathodes at long life.

The writer wants to thank Mr. S. Costa and Mr. H. Benda, who have done most of the experimental work.

[10] Described by A. L. Reimann, *Thermionic Emission* (Chapman and Hall, Ltd., London, 1934), p. 169.

Modes in Wave Guides Containing Ferrites*

M. L. KALES
Naval Research Laboratory, Washington D. C.
(Received December 8, 1952)

When a static magnetic field is applied in the longitudinal direction to a wave guide containing a ferrite material, the medium within the guide becomes anisotropic. It has been shown by Polder that when a microwave field is propagating within such a medium, the relation between the components of **B** and **H** is given by: $B_x=\mu H_x-j\mu' H_y$, $B_y=j\mu' H_x+\mu H_y$, $B_z=H_z$, where μ and μ' depend on the static magnetic field. Using these equations, an investigation of the wave-guide modes is made. Formal solutions are obtained for guides having circular symmetry. These solutions depend on the solution of some rather involved eigenvalue equations which determine the propagation constants, and which probably can be solved only by numerical methods.

INTRODUCTION

RECENT researches with ferrite materials[1–5] demonstrated the possibility of producing large changes in phase or rotation of polarization for electromagnetic waves at microwave frequencies. Normally, when these waves pass through such materials, the materials behave essentially like isotropic dielectrics. If, however, a static magnetic field is impressed upon the ferrite, it becomes anisotropic in its magnetic properties. As a consequence of this anisotropy, either the phase or the polarization of the microwave field upon passing through the ferrite can be made to vary with the static magnetic field. Since this field can be made to vary by purely electrical means, it is possible, with the aid of the ferrite materials, to construct microwave components that will rapidly vary either phase or polarization electrically.

The results obtained in a theoretical investigation of the modes in wave guides containing ferrites represent only a first step towards a more complete understanding of their behavior. Much more work, both theoretical and experimental, remains to be done.

GENERAL RELATIONS

Polder[1] demonstrated that if a static magnetic field is established in a ferrite medium in the direction of the z axis and if a microwave field is propagating in the ferrite medium, the components of **B** and **H** of the electromagnetic field are related by equations of the following form:

$$\begin{aligned} B_x&=\mu H_x-j\mu' H_y \quad [j=(-1)^{\frac{1}{2}}],\\ B_y&=j\mu' H_x+\mu H_y,\\ B_z&=\mu_z H_z. \end{aligned} \tag{1}$$

Formulas for μ, μ', and μ_z were obtained by Polder for the case in which the static magnetic field $\mathbf{H}_c$ saturates the ferrite and the magnitude of the microwave **H** is small compared to $\mathbf{H}_c$. Polder pointed out, however, that if the ferrite is uniformly magnetized in the direction of the z axis, Eqs. (1) will represent the most general linear relationship between **B** and **H**. This hypothesis will be assumed for the remainder of the discussion.

Consider now a cylindrical wave guide (a) with its axis in the direction of the z axis and (b) filled with ferrite material uniformly magnetized in the direction of the z axis. It will be assumed that the medium is isotropic and homogeneous with respect to its dielectric properties and that μ, μ', and μ_z are independent of position. Let $\mathbf{i}_z$ denote the unit vector in the direction of the z axis. The subscript t will be used to denote the projection of a vector on the $x-y$ plane. The relations between **B** and **H** and **D** and **E** may now be expressed in vector form by

$$\mathbf{B}=\mu\mathbf{H}_t+j\mu'\mathbf{i}_z\times\mathbf{H}_t+\mu_z H_z\mathbf{i}_z$$

and

$$\mathbf{D}=\epsilon\mathbf{E}.$$

Assuming fields which have an $e^{j\omega t}$ time dependence and an $e^{\gamma z}z$ dependence, the space-dependent parts of the field will be represented by

$$\mathbf{E}(x,y,z)=\mathbf{E}(x,y)e^{\gamma z},\quad \mathbf{H}(x,y,z)=\mathbf{H}(x,y)e^{\gamma z},\quad \mathbf{B}(x,y,z)=\mathbf{B}(x,y)e^{\gamma z}.$$

In the remaining part of the discussion the vectors **E**, **H**, and **B** will denote the vector functions of x and y only.

From Maxwell's equations,

$$\nabla_t\times\mathbf{E}+\gamma\mathbf{i}_z\times\mathbf{E}=-j\omega\mathbf{B}$$

and

$$\nabla_t\times\mathbf{H}+\gamma\mathbf{i}_z\times\mathbf{H}=j\omega\epsilon\mathbf{E},$$

but

$$\nabla_t\times\mathbf{E}=\nabla_t\times\mathbf{E}_t+\nabla_t E_z\times\mathbf{i}_z$$

and

$$\nabla_t\times\mathbf{H}=\nabla_t\times\mathbf{H}_t+\nabla_t H_z\times\mathbf{i}_z.$$

Substituting the last pair of equations into the preceding

---

* While the material contained in the present paper was being prepared for Naval Research Laboratory Report No. 4027, August 8, 1952, a letter to the editor by H. Suhl and L. R. Walker appeared on the same subject in the April 1, 1952, issue of *The Physical Review*. From this letter it appears that results similar to those of the present paper were obtained.

[1] D. Polder, Phil. Mag. **40**, 99 (1949).
[2] F. F. Roberts, J. phys. et radium, **12**, 305 (1951).
[3] C. L. Hogan, Bell System Tech. J. **31**, 1 (1952).
[4] C. H. Luhrs, Proc. Inst. Radio Engrs. **40**, 76 (1952).
[5] Sakiotis, Simmons, and Chait, "Applications of Ferrites to the Microwave Antenna Problem," Reprint No. 28–52 from NRL Progress Report (January 1952).

pair we get

$$\nabla_t \times \mathbf{E}_t + \mathbf{i}_z \times (\gamma \mathbf{E}_t - \nabla_t E_z) = -j\omega \mathbf{B}$$

and

$$\nabla_t \times \mathbf{H}_t + \mathbf{i}_z \times (\gamma \mathbf{H}_t - \nabla_t H_z) = j\omega \epsilon \mathbf{E}.$$

Equating the longitudinal and transverse components, respectively, gives

$$\nabla_t \times \mathbf{E}_t = -j\omega \mu_z H_z \mathbf{i}_z, \quad (2a)$$

$$\nabla_t \times \mathbf{H}_t = j\omega \epsilon E_z \mathbf{i}_z, \quad (2b)$$

$$\mathbf{i}_z \times \gamma \mathbf{E}_t + j\omega \mu \mathbf{H}_t - \omega \mu' \mathbf{i}_z \times \mathbf{H}_t = \mathbf{i}_z \times \nabla_t E_z, \quad (2c)$$

$$\mathbf{i}_z \times \gamma \mathbf{H}_t - j\omega \epsilon \mathbf{E}_t = \mathbf{i}_z \times \nabla_t H_z. \quad (2d)$$

If we let $k^2 = \omega^2 \epsilon \mu$, $k'^2 = \omega^2 \epsilon \mu'$, and $K^2 = \omega^2 \epsilon \mu + \gamma^2$, then from Eqs. (2c) and (2d) we find that

$$K^4 \mathbf{E}_t + jK^2 k'^2 \mathbf{i}_z \times \mathbf{E}_t = K^2 \{\gamma \nabla_t E_z + j\omega \mu \mathbf{i}_z \times \nabla_t H_z + \omega \mu' \nabla_t H_z\}, \quad (3a)$$

$$k'^4 \mathbf{E}_t + jK^2 k'^2 \mathbf{i}_z \times \mathbf{E}_t = jk'^2 \{\gamma \mathbf{i}_z \times \nabla_t E_z - j\omega \mu \nabla_t H_z + \omega \mu' \mathbf{i}_z \times \nabla_t H_z\}. \quad (3b)$$

The last pair of equations and a similar pair which can be obtained for $\mathbf{H}_t$ can be solved for $\mathbf{E}_t$ and $\mathbf{H}_t$, giving

$$(K^4 - k'^4)\mathbf{E}_t = \nabla_t (K^2 \gamma E_z + \omega \gamma^2 \mu' H_z) + j\mathbf{i}_z \times \nabla_t \{\omega (K^2 \mu - k'^2 \mu') H_z - \gamma k'^2 E_z\}, \quad (4a)$$

$$(K^4 - k'^4)\mathbf{H}_t = \nabla_t (K^2 \gamma H_z + \omega \epsilon k'^2 E_z) - j\mathbf{i}_z \times \nabla_t (K^2 \omega \epsilon E_z + \gamma k'^2 H_z). \quad (4b)$$

From the divergence equations $\nabla \cdot \mathbf{E} e^{\gamma z} = 0$ and $\nabla \cdot \mathbf{B} e^{\gamma z} = 0$ we find

$$\nabla_t \cdot \mathbf{E}_t = -\gamma E_z$$

and

$$\nabla_t \cdot \mathbf{B}_t = -\gamma B_z. \quad (5)$$

Expressing $\mathbf{B}$ in terms of $\mathbf{H}$, the last equation leads to

$$\nabla_t \cdot \mathbf{H}_t = -\gamma \frac{\mu_z}{\mu} H_z - \omega \epsilon \frac{\mu'}{\mu} E_z.$$

Taking the divergence of both sides of Eqs. (4) and combining the results with Eqs. (5) leads to

$$\nabla_t^2 E_z + aE_z + bH_z = 0$$

and

$$\nabla_t^2 H_z + cH_z + dE_z = 0, \quad (6)$$

where $a = K^2 - k'^2 \mu'/\mu$, $b = -\omega \gamma \mu_z \mu'/\mu$, $c = K^2 \mu_z/\mu$, $d = \omega \gamma \epsilon \mu'/\mu$. (A list of special symbols appears in the appendix.)

The boundary conditions needed for solving Eqs. (6) are obtained by equating to zero the tangential components of $\mathbf{E}$ on the surface of the wave guide. If the boundary of the cross section of the guide is denoted by C, then the tangential components of $\mathbf{E}$ are E_z and the component of $\mathbf{E}_t$ which is tangent to C. If $K^4 - k'^4 \neq 0$, then the latter may be obtained from Eq. (4a). Thus the boundary conditions are

$$E_z = 0$$

and

$$jk'^2 \gamma \frac{\partial E_z}{\partial n} + \omega \gamma^2 \mu' \frac{\partial H_z}{\partial \tau} - j\omega (K^2 \mu - k'^2 \mu') \frac{\partial H_z}{\partial n} = 0 \Bigg\} \text{on } C,$$

where $\partial/\partial n$ and $\partial/\partial \tau$ represent the normal and tangential derivatives on C, respectively.

NONEXISTENCE OF *TE*, *TM*, OR *TEM* MODES

Let us first consider the possibility of having modes for which E_z or H_z is identically zero. We shall assume first that $\gamma \neq 0$ and hence that b and d in Eqs. (6) are both different from zero. It follows at once from the same equations that neither E_z nor H_z can be identically zero unless both are zero. If both E_z and H_z vanish, we see from Eq. (4a) that either $\mathbf{E}_t = 0$ or $K^4 - k'^4 = 0$. If $\mathbf{E}_t = 0$, then the field vanishes identically. Suppose, then, that $K^4 - k'^4 = 0$. It then follows from Eq. (3a) that

$$\mathbf{E}_t = \pm j\mathbf{i}_z \times \mathbf{E}_t. \quad (7)$$

From Eq. (2a) it follows that when $H_z \equiv 0$,

$$\nabla_t \times \mathbf{E}_t \equiv 0,$$

Hence a function $\psi(x, y)$ exists such that

$$\mathbf{E}_t = \nabla_t \psi. \quad (8)$$

From Eqs. (7) and (8), then

$$\nabla_t \psi = \pm j\mathbf{i}_z \times \nabla_t \psi. \quad (9)$$

If we now let $\psi \equiv u(x, y) \pm jv(x, y)$ where u and v are real functions, we see that Eq. (9) is simply the vector form of the Cauchy-Riemann equations which express the fact that ψ is an analytic function of the complex variable $\zeta = x + jy$. Since the tangential component of $\mathbf{E}_t$ must vanish on C, we see that

$$|(E_t)_{\tan}| = \left|\frac{\partial \psi}{\partial \tau}\right| = |\psi'(\zeta)| = 0 \text{ on } C.$$

Since $\psi'(\zeta) = 0$ on C, it follows from function theory that $\psi(\zeta)$ is constant within the region bounded by C. Hence,

$$\mathbf{E}_t = \nabla_t \psi = 0.$$

Thus, if $\gamma \neq 0$, no modes are possible for which either E_z or H_z is identically zero. In other words, no *TE*, *TM*, or *TEM* modes exist.

CUT-OFF FREQUENCIES

Let us now consider the case of $\gamma = 0$. Here Eqs. (6) take the form

$$\nabla_t^2 E_z + aE_z = 0$$

and

$$\nabla_t^2 H_z + cH_z = 0, \quad (10)$$

and the boundary condition becomes

$$\left.\begin{aligned} E_z&=0 \\ \frac{\partial H_z}{\partial n}&=0 \end{aligned}\right\} \text{on } C.$$

Thus, it is seen that when $\gamma=0$ the solutions can be separated into TE and TM modes as in the case of a homogeneous isotropic medium. If we define the cut-off frequencies as those for which the wavelength becomes infinite, or in other words, for which $\gamma=0$, then these cut-off frequencies can be determined from the allowed values of a and c. Let K_c^2 denote the values of a or c for which solutions exist to Eqs. (10) that satisfy the boundary conditions. Since when $\gamma=0$, $a=\omega^2\epsilon\mu - \omega^2\epsilon\mu'^2/\mu$ and $c=\omega^2\epsilon\mu_z$, we find that

$$\omega_c=\frac{K_c}{(\epsilon\mu)^{\frac{1}{2}}\left[1-\left(\frac{\mu'}{\mu}\right)^2\right]^{\frac{1}{2}}}=\frac{\omega_c^{(0)}}{\left[1-\left(\frac{\mu'}{\mu}\right)^2\right]^{\frac{1}{2}}} \text{ for } TM \text{ modes}$$

and

$$\omega_c=\frac{K_c}{(\epsilon\mu_z)^{\frac{1}{2}}}=\left(\frac{\mu}{\mu_z}\right)^{\frac{1}{2}}\omega_c^{(0)} \text{ for } TE \text{ modes,}$$

where $\omega_c^{(0)}$ denotes the cut-off frequencies in a homogeneous isotropic medium of permeability μ.

TRANSFORMATION TO A PAIR OF WAVE EQUATIONS

Returning to Eqs. (6), let us attempt to transform them to a pair of wave equations by expressing E_z and H_z in terms of two new variables u_1 and u_2 as follows:

$$E_z=p_1u_1+p_2u_2$$

and

$$H_z=q_1u_1+q_2u_2,$$

where p_1, p_2, q_1, and q_2 are constants to be determined. In order not to lose generality we impose the condition

$$\begin{vmatrix} p_1 & p_2 \\ q_1 & q_2 \end{vmatrix}\neq 0. \tag{11}$$

Substituting for E_z and H_z in Eqs. (6) we get

$$\begin{aligned} &p_1\nabla_t^2u_1+(ap_1+bq_1)u_1+p_2\nabla_t^2u_2+(ap_2+bq_2)u_2=0, \\ &q_1\nabla_t^2u_1+(cq_1+dp_1)u_1+q_2\nabla_t^2u_2+(cq_2+dp_2)u_2=0. \end{aligned} \tag{12}$$

Now let

$$\begin{aligned} ap_1+bq_1&=s_1p_1 \quad ap_2+bq_2=s_2p_2 \\ dp_1+cq_1&=s_1q_1 \quad dp_2+cq_2=s_2q_2, \end{aligned} \tag{13}$$

where s_1 and s_2 are two more constants to be determined. Equations (12) then take the form

$$\begin{aligned} &p_1(\nabla_t^2u_1+s_1u_1)+p_2(\nabla_t^2u_2+s_2u_2)=0 \\ &\text{and} \\ &q_1(\nabla_t^2u_1+s_1u_1)+q_2(\nabla_t^2u_2+s_2u_2)=0. \end{aligned} \tag{14}$$

If Condition (11) is satisfied, it follows from Eq. (14) that

$$\nabla_t^2u_1+s_1u_1=0$$

and

$$\nabla_t^2u_2+s_2u_2=0. \tag{15}$$

In order that Eqs. (13) have a nontrivial solution, we must have

$$\begin{vmatrix} a-s_1 & b \\ d & c-s_1 \end{vmatrix}=0, \quad \begin{vmatrix} a-s_2 & b \\ d & c-s_2 \end{vmatrix}=0. \tag{16}$$

Equations (16) and (11) will be satisfied if we take for s_1 and s_2 the two (in general) distinct roots of the quadratic in s,

$$\begin{vmatrix} a-s & b \\ d & c-s \end{vmatrix}=0.$$

For convenience we can let

$$p_1=s_1, \quad p_2=s_2.$$

Then

$$q_1=\frac{(s_1-a)s_1}{b}, \quad q_2=\frac{(s_2-a)s_2}{b}.$$

If we now express the various quantities appearing in Eqs. (4) in terms of the new variables, we find that

$$K^4-k'^4=\frac{\mu}{\mu_z}s_1s_2,$$

$$K^2\gamma E_z+\omega\gamma^2\mu' H_z=\gamma\frac{\mu}{\mu_z}s_1s_2(u_1+u_2),$$

$$\omega(K^2\mu-k'^2\mu')H_z-\gamma k'^2E_z=\omega\mu s_1s_2\left(\frac{s_1-a}{b}u_1+\frac{s_2-a}{b}u_2\right),$$

$$K^2\gamma H_z+\omega\epsilon k'^2E_z=\gamma s_1s_2\left(\frac{s_1-K^2}{b}u_1+\frac{s_2-K^2}{b}u_2\right),$$

$$K^2\omega\epsilon E_z+\gamma k'^2H_z=\omega\epsilon\frac{\mu}{\mu_z}s_1s_2(u_1+u_2).$$

Substituting these results in Eqs. (4) and combining them with results previously obtained, leads to the following solution.

Let s_1 and s_2 be the two roots of the quadratic equation

$$\begin{vmatrix} a-s & b \\ d & c-s \end{vmatrix}=0.$$

Let u_1 and u_2 be solutions of the wave equations

$$\nabla_t^2u_1+s_1u_1=0$$

and

$$\nabla_t^2u_2+s_2u_2=0,$$

subject to the boundary conditions

$$\left.\begin{aligned} &s_1u_1+s_2u_2=0 \\ \text{and} \\ &\gamma\frac{\partial}{\partial\tau}(u_1+u_2)-\frac{j\mu}{\gamma\mu'}\frac{\partial}{\partial n}\{(s_1-a)u_1+(s_2-a)u_2\}=0 \end{aligned}\right\}\text{on } C.$$

Then the field inside the guide is given by the following equations:

$$E_z=s_1u_1+s_2u_2, \tag{17a}$$

$$H_z=\frac{(s_1-a)s_1}{b}u_1+\frac{(s_2-a)s_2}{b}u_2, \tag{17b}$$

$$\mathbf{E}_t=\gamma\nabla_t(u_1+u_2)-\frac{j\mu}{\gamma\mu'}\mathbf{i}_z\times\nabla_t\{(s_1-a)u_1+(s_2-a)u_2\}, \tag{17c}$$

$$\mathbf{H}_t=\frac{1}{\omega\mu'}\nabla_t\{(K^2-s_1)u_1+(K^2-s_2)u_2\}-j\omega\epsilon\mathbf{i}_z\times\nabla_t(u_1+u_2). \tag{17d}$$

CHARACTERISTIC EQUATION FOR THE FILLED ROUND GUIDE

Up to this point the cross section of the wave guide has been assumed to be arbitrary. To carry the solution a step further, we shall now restrict ourselves to the case of a circular cross section. Let the radius be r_0, and assume solutions which have an $e^{jn\theta}$ θ-dependence. Then

$$\left.\begin{aligned} &u_1=A_1J_n(s_1^{\frac{1}{2}}r)e^{jn\theta} \\ \text{and} \qquad &n=0,\pm1,\pm2,\cdots, \\ &u_2=A_2J_n(s_2^{\frac{1}{2}}r)e^{jn\theta} \end{aligned}\right.$$

where $J_n(x)$ denotes the Bessel function of order n and A_1 and A_2 are constants. Imposing the boundary conditions, we get

$$s_1J_n(s_1^{\frac{1}{2}}r_0)A_1+s_2J_n(s_2^{\frac{1}{2}}r_0)A_2=0$$

and

$$\frac{n\gamma}{r_0}\{J_n(s_1^{\frac{1}{2}}r_0)A_1+J_n(s_2^{\frac{1}{2}}r_0)A_2\}-\frac{\mu}{\gamma\mu'}\times\{(s_1-a)s_1^{\frac{1}{2}}J_n'(s_1^{\frac{1}{2}}r_0)A_1+(s_2-a)s_2^{\frac{1}{2}}J_n'(s_2^{\frac{1}{2}}r_0)A_2\}=0.$$

For a nontrivial solution we require that the determinant of the coefficients of A_1 and A_2 in the last pair of equations be zero. This leads to the equation

$$\mu\left\{\left(1-\frac{a}{s_1}\right)(s_1r_0)^{\frac{1}{2}}\frac{J_n'[(s_1r_0)^{\frac{1}{2}}]}{J_n[(s_1r_0)^{\frac{1}{2}}]}-\left(1-\frac{a}{s_2}\right)(s_2r_0)^{\frac{1}{2}}\frac{J_n'(s_2r_0)^{\frac{1}{2}}}{J_n(s_2r_0)^{\frac{1}{2}}}\right\}+n\mu'\gamma^2\left(\frac{1}{s_1}-\frac{1}{s_2}\right)=0. \tag{18}$$

The quadratic equation in s whose two roots are s_1 and s_2 may be written as follows:

$$s^2-\left(K^2+K^2\frac{\mu_z}{\mu}-k'^2\frac{\mu'}{\mu}\right)s+\frac{\mu_z}{\mu}(K^4-k'^4)=0.$$

Since $K^2=k^2+\gamma^2$, we see that s_1 and s_2 are functions of γ^2. In addition, $a=K^2-(k'^2\mu'/\mu)$ is a function of γ^2. Thus the left-hand member of Eq. (18) may be regarded as an analytic function of γ^2 whose zeros determine the propagation constants. Alternatively, γ^2 may be eliminated and a pair of simultaneous equations in s_1 and s_2 obtained. Once s_1 and s_2 are known, γ^2 as well as the functions u_1 and u_2 may be determined. To see how this can be done, we observe that from the quadratic equation which determines s_1 and s_2

$$s_1+s_2=a+c=K^2\left(1+\frac{\mu_z}{\mu}\right)-k'^2\frac{\mu'}{\mu}$$

and

$$s_1s_2=ac-bd=\mu_z/\mu(K^4-k'^4).$$

Eliminating γ^2, we find

$$\gamma^2=\frac{\mu(s_1+s_2)+k'^2\mu'-k^2(\mu+\mu_z)}{\mu+\mu_z} \tag{19}$$

and

$$\left\{\frac{\mu(s_1+s_2)+k'^2\mu'}{\mu+\mu_z}\right\}^2-\frac{\mu}{\mu_z}s_1s_2-k'^4=0. \tag{20}$$

If we now eliminate γ^2 from Eq. (18) with the aid of Eq. (19), then Eqs. (18) and (20) become a pair of simultaneous equations for the determination of s_1 and s_2.

Before concluding this section let us consider the special case $n=\pm1$. On the axis of the guide

$$\nabla_tJ_1(s^{\frac{1}{2}}r)e^{\pm j\theta}]_{r=0}=\left\{s^{\frac{1}{2}}J_1'(s^{\frac{1}{2}}r)\mathbf{i}_r+j\frac{J_1(s^{\frac{1}{2}}r)}{r}\mathbf{i}_\theta\right\}e^{\pm j\theta}\Big]_{r=0}$$

$$=\frac{s^{\frac{1}{2}}}{2}(\mathbf{i}_r\pm j\mathbf{i}_\theta)e^{\pm j\theta}, \tag{21}$$

where $\mathbf{i}_r$ and $\mathbf{i}_\theta$ denote unit vectors in the directions of increasing r and increasing θ. From Eqs. (17c), (17d), and (21) it follows that for $n=\pm1$, $\mathbf{E}_t$ and $\mathbf{H}_t$ represent right- or left-hand circularly polarized vectors on the axis of the guide. However, excluding the exceptional cases where $\gamma=0$ or $s_1=s_2$, the propagation constants for $n=1$, or -1 will be different. This can be seen from Eq. (18). If we assume that a common solution γ exists for $n=\pm1$, substituting $n=\pm1$ in Eq. (18) and subtracting would lead to the excluded equation

$$2\mu'\gamma^2\left(\frac{1}{s_1}-\frac{1}{s_2}\right)=0.$$

Thus, if two modes for which $n=\pm 1$ are propagating in the same direction along the guide, the corresponding $\mathbf{E}_t$ and $\mathbf{H}_t$ vectors will be rotating in opposite directions on the axis. If the amplitudes of $\mathbf{E}_t$ for the two modes are equal on axis, the resultant vector sum will be linearly polarized. Since the velocities of propagation of the two modes are different, the relative phase between the two circularly polarized components will vary with position along the z axis; hence the orientation of the linearly polarized $\mathbf{E}_t$ vector will vary with position along the z axis. The same will be true for $\mathbf{H}_t$. This is the well-known "Faraday Effect."

CHARACTERISTIC EQUATION FOR THE PARTIALLY FILLED ROUND GUIDE

In many applications the ferrite material only partially fills the wave guide. In this case the equation which determines the propagation constants becomes rather formidable. For the sake of completeness, however, the derivation of this equation will be outlined briefly.

Let the region $0\leqq r\leqq r_1$ contain the ferrite, and let the region $r_1\leqq r\leqq r_0$ be filled with a homogeneous isotropic dielectric. Quantities associated with the latter region will be denoted by the subscript 1. Assuming again solutions with an $e^{in\theta}$ θ-dependence, we find for the portion of the cross section which does not contain the ferrite

$$E_{1z}=A_3\left\{\frac{J_n(K_1r)}{J_n(K_1r_0)}-\frac{Y_n(K_1r)}{Y_n(K_1r_0)}\right\}e^{in\theta}\equiv A_3F(r)e^{in\theta}$$

and

$$H_{1z}=A_4\left\{\frac{J_n(K_1r)}{J_n'(K_1r_0)}-\frac{Y_n(K_1r)}{Y_n'(K_1r_0)}\right\}e^{in\theta}\equiv A_4G(r)e^{in\theta},$$

where $Y_n(x)$ is the Neumann function of order n, $K_1^2=\omega^2\epsilon_1\mu_1+\gamma^2$, and the prime denotes differentiation with respect to the argument. In the ferrite region

$$E_z=\{s_1J_n(s_1^{\frac{1}{2}}r)A_1+s_2J_n(s_2^{\frac{1}{2}}r)A_2\}e^{in\theta}$$

and

$$H_z=\left\{\frac{(s_1-a)}{b}s_1J_n(s_1^{\frac{1}{2}}r)A_1+\frac{(s_2-a)s_2}{b}J_n(s_2^{\frac{1}{2}}r)A_2\right\}e^{in\theta}.$$

Note that γ is the same for both regions. From the expressions for E_z and H_z in both regions, the transverse components of $\mathbf{E}$ and $\mathbf{H}$ which are tangent to the circle $r=r_1$ can be obtained. Imposing the condition that the tangential components of $\mathbf{E}$ and $\mathbf{H}$ are continuous across the boundary $r=r_1$ leads to a set of four linear homogeneous equations in the constants A_1, A_2, A_3, and A_4. The characteristic equation for determining the possible values of γ is then obtained by equating to zero the determinant of this system of equations. This leads to the equation

$$P_1Q_2-P_2Q_1=0, \tag{22}$$

where for $i=1, 2$

$$P_i=\{n\gamma^2\mu'\mu_z(K_1^2-s_i)G(r_1)+\mu_1\mu r_1G'(r_1)s_i(s_i-a)\}J_n(s_i^{\frac{1}{2}}r_1)-\mu\mu_zK_1^2G(r_1)(s_i-a)s_i^{\frac{1}{2}}r_1J_n'(s_i^{\frac{1}{2}}r_1),$$

$$Q_i=\{n(K^2-s_i)K_1^2F(r_1)\mu_z+n\mu s_i(s_i-a)F(r_1)+\omega^2\epsilon_1\mu'\mu_zr_1F'(r_1)s_i\}J_n(s_i^{\frac{1}{2}}r_1)-\omega^2\epsilon\mu'\mu_zK_1^2F(r_1)s_i^{\frac{1}{2}}r_1J_n'(s_i^{\frac{1}{2}}r_1).$$

CONCLUSION

Equations (18) and (22) determine the possible values of γ for the case of the wave guide of circular cross section which is filled or partially filled with ferrite. Once the values of γ are found, the corresponding modes can be obtained from the appropriate equations which have been given. It does not appear likely that Eqs. (18) and (22) can be solved explicitly for γ as a function of the various parameters involved. For given values of the parameters, however, solutions can probably be obtained by numerical methods of calculation.

APPENDIX. SPECIAL SYMBOLS

$k^2=\omega^2\epsilon\mu$
$k'^2=\omega^2\epsilon\mu'$
$K^2=\omega^2\epsilon\mu+\gamma^2$
$a=K^2-k'^2\mu'/\mu$
$b=-\omega\gamma\mu_z\mu'/\mu$
$c=K^2\mu_z/\mu$
$d=\omega\gamma\epsilon\mu'/\mu$
s_1, s_2 = the two roots of the quadratic equation in s
$s^2-(a+c)s+ac-bd=0.$

JOURNAL OF APPLIED PHYSICS VOLUME 24, NUMBER 5 MAY, 1953

Response of Linear Time-Dependent Systems to Random Inputs

D. B. DUNCAN
Electro-Mechanical Engineering Department, North American Aviation, Inc., Downey, California
(Received December 15, 1952)

A nonstatistical differential equation is derived for the mean square response of linear systems to random inputs. This equation can be used to determine analytically the steady state and transient response of systems with constant coefficients. It is also applicable to time dependent systems. While the solution cannot in general be obtained analytically in this case, the equation is of the type that can be immediately mechanized on analog computers.

INTRODUCTION

TECHNIQUES are now known which can be used to calculate the transient and steady-state values of the mean square output of a constant parameter linear system excited by noise. The results can be expressed either as integrals in the frequency domain of the power spectrum of the input and the transfer function of the system or as integrals in the time domain of the autocorrelation function of the input and the weighting function of the system. These results can be generalized to time dependent systems without formal difficulty, but the resulting integrals cannot in general be evaluated analytically and are in a cumbersome form for numerical evaluation. One method that has been used when it is necessary to obtain numerical results for the time dependent case is to simulate the system with an analog computer and excite it with a noise source. This approach has the serious disadvantage of requiring a large amount of data to obtain reliable results.

In this paper another approach to the problem of determining the mean square output of a linear system is developed. A single differential equation for this quantity is derived; the driving function for this equation is nonstatistical in nature. For constant parameter systems, this equation can be solved to give an analytic expression for either the transient or steady-state value. It appears that this process involves about the same number of mathematical steps as the standard techniques. For time-dependent systems, however, it appears to offer definite advantages. The problem is reduced to that of finding the solution of a linear time-dependent differential equation. Although an analytic solution cannot in general be obtained, the well-developed theory of differential equations can be used for results of a general nature and computational methods, such as the use of an analog computer, can be employed directly to obtain a numerical solution.

In the first section, this differential equation is derived with the restriction that white noise is exciting the system. It is then shown that the general case where the noise has a known autocorrelation function or power spectrum can be reduced to that form.

SYSTEM EXCITED BY WHITE NOISE

The equation describing a linear time-dependent system excited by white noise is

$$A_n(t)\frac{d^n x}{dt^n}+A_{n-1}(t)\frac{d^{n-1}x}{dt^{n-1}}+\cdots+A_0(t)x=f(t), \quad (1)$$

where the coefficients A_i are known functions of time and $f(t)$ is a random variable with an autocorrelation function $\phi(t)=\delta(t)$. It will be shown later that this restriction on the form of the autocorrelation function can be made without loss of generality.

The response of the system described by Eq. (1) can be obtained formally and is, for zero initial conditions,

$$x(t)=\int_0^t g(t,u)f(t-u)du, \quad (2)$$

where $g(t,u)$ is the response at time t to a unit impulse at time $t-u$. An expression for the transient value of the mean square error which can be obtained immediately is

$$\langle x^2(t)\rangle_{av}=\int_0^t\int_0^t g(t,u)g(t,v)\phi(u-v)dudv. \quad (3)$$

While this equation gives the formal solution, its practical usefulness is limited. In general, $g(t,u)$ must be obtained as a function of the two variables by computer techniques. It is then necessary to carry out the integration numerically. The problem is particularly unsuited for solution with an analog computer.

An alternate method of obtaining $\langle x^2(t)\rangle_{av}$ is the following: Form the set of equations

$$\frac{d}{dt}\left[\frac{d^i x}{dt^i}\frac{d^j x}{dt^j}\right]=\frac{d^{i+1}x}{dt^{i+1}}\frac{d^j x}{dt^j}+\frac{d^i x}{dt^i}\frac{d^{j+1}x}{dt^{j+1}}$$

$$i,j=0,1,\cdots,n-1; \quad (4)$$

for i and j less than $n-1$, terms of the same form result from the differentiation process. For i or j equal to $n-1$, a term involving the nth derivative is obtained. However, from Eq. (1), this can be expressed as a function of the lower order derivatives. The process will be to substitute $A_n^{-1}[f-A_{n-1}d^{n-1}x/dt^{n-1}-\cdots-A_0x]$, wher-

ever d^nx/dt^n occurs and to take the ensemble average of the resulting equations.

In this process, it will be necessary to calculate ensemble averages of terms of the form fd^kx/dt^k, where k has values from 0 to $n-1$. Analogous to Eq. (2), the kth derivative of x is given by

$$\frac{d^kx}{dt^k}=\int_0^t g_k(t,u)f(t-u)du, \tag{5}$$

where $g_k(t,u)$ is the response of d^kx/dt^k at time t to a unit impulse at time $t-u$. Multiplying by $f(t)$ and averaging, one obtains

$$\left\langle f\frac{d^kx}{dt^k}\right\rangle_{Av}=\int_0^t g_k(t,u)\delta(u)du. \tag{6}$$

Since this equation involves a delta-function at the origin, this average value depends only on $g_k(t,0)$. Since this function does not depend on past history, it is the same for the system under consideration as for a system with constant coefficients having the values A_i at time t and is equal to A_n^{-1} for k equal to $n-1$ and zero otherwise. This same result can be obtained, perhaps with more rigor, from the mathematical definition of $g_k(t,u)$ as a solution of the homogenous equation with certain discontinuities in its derivatives at $u=0$.[1] Carrying out the integration, a factor of $\frac{1}{2}$ is introduced because of the delta-function at the end of the range of integration, and one obtains

$$\left\langle f\frac{d^kx}{dt^k}\right\rangle_{Av}=\tfrac{1}{2}A_n^{-1}\delta_{n-1}{}^k, \tag{7}$$

where $\delta_{n-1}{}^k$ is the Kronecker delta. Define the variables

$$\theta_{ij}=\left\langle\frac{d^ix}{dt^i}\frac{d^jx}{dt^j}\right\rangle_{Av},$$

where i and j have values from 0 to $n-1$. Note that $\theta_{ij}=\theta_{ji}$, so that there are $\frac{1}{2}n(n+1)$ independent variables.

Forming the ensemble average of Eq. (4), one obtains

$$\begin{aligned}
&\frac{d\theta_{ij}}{dt}=\theta_{i+1,\,j}+\theta_{i,\,j+1}\quad i,j=0,1,\cdots,n-2,\\
&A_n\frac{d\theta_{i,\,n-1}}{dt}=A_n\theta_{i+1,\,n-1}-[A_{n-1}\theta_{i,\,n-1}+A_{n-2}\theta_{i,\,n-2}\\
&\qquad+\cdots+A_0\theta_{i,\,0}]\quad i=0,1,n-2,\\
&A_n{}^2\frac{d\theta_{n-1,\,n-1}}{dt}=1-2A_n[A_{n-1}\theta_{n-1,\,n-1}\\
&\qquad+A_{n-2}\theta_{n-1,\,n-2}+\cdots+A_0\theta_{n-1,\,0}].
\end{aligned} \tag{8}$$

These are now a set of $\frac{1}{2}n(n+1)$ equations in $\frac{1}{2}n(n+1)$ independent variables and serve to determine the solution for $\langle x^2\rangle_{Av}=\theta_{00}$, as well as other stochastic variables.

REDUCTION OF EQUATIONS TO THOSE OF A SYSTEM EXCITED BY WHITE NOISE

The solution outlined above depends upon the random variable having an autocorrelation function which is a delta-function. By the artifice of passing white noise through the appropriate filter, the results can be made applicable to the general case. The power spectrum $\Phi(j\omega)$, which is the Fourier transform of $\phi(t)$, can be factored into $\Psi(j\omega)\Psi(-j\omega)$.[2] If white noise of unit power per unit frequency is introduced into a system whose transfer function is $[\Psi(s)]^{-1}$, with initial conditions corresponding to steady-state values, the output will be a random variable of an autocorrelation function $\phi(t)$. Thus, one can consider white noise introduced into an over-all system consistlng of a linear constant coefficient system followed by the variable system under consideration. Mathematically, this result can be obtained by operating on Eq. (1) with the differential operator $[\Psi(p)]^{-1}$ and identifying $[\Psi(p)]^{-1}f(t)$ as white noise. A higher order equation [with the same form as Eq. (1)] is obtained by this process with coefficients which are functions of the coefficients A_i and their derivatives. It must be noted that the higher order initial conditions of this new equation are no longer arbitrary but are determined from the relationships

$$\frac{d^i}{dt^i}\left[A_n\frac{d^nx}{dt^n}\right]_{t=0}=\frac{d^if}{dt^i}\Big]_{t=0}. \tag{9}$$

This process, therefore, permits the results of the previous paragraph to be applied when the random variable has a general autocorrelation function. The initial conditions implied in Eq. (9) must be used in the solution.

DISCUSSION

For the case of constant coefficients, Eq. (8) can be solved by analytical methods and the result obtained in closed form. Both the steady state and the transient values can be found. The important feature of the process, however, is that it reduces the problem, even for the time variable system, to the solution of a nonstatistical differential equation. Although the equation will generally not be solvable in closed form, the powerful and well-developed theory of ordinary differential equations can be used for results of a general nature. Furthermore, the form of the equation is ideal for mechanization on an analog computer such as the REAC.

[1] E. L. Ince, *Ordinary Differential Equations* (Dover Publications, Inc., New York, 1926).

[2] N. Wiener, *Extrapolation, Interpolation and Smoothing of Stationary Time Series* (John Wiley and Sons, Inc., New York, 1949).

ILLUSTRATIVE EXAMPLE

As an example, consider the solution of the following simple equation with constant coefficients:

$$\frac{dx}{dt}+ax=f(t), \tag{10}$$

subject to the initial condition $x(0)=0$, where $f(t)$ is a random variable with autocorrelation function, $\phi_0\exp(-b|t|)$. For this autocorrelation function, $[\Psi(p)]^{-1}=(2\phi_0 b)^{-\frac{1}{2}}(p+b)$, so the solution of Eq. (10) reduces to the solution of the following equation:

$$\frac{d^2x}{dt^2}+(a+b)\frac{dx}{dt}+abx=(2\phi b)^{\frac{1}{2}}g(t), \tag{11}$$

subject to the initial conditions $x(0)=0$, $dx(0)/dt=f(0)$ and where $g(t)$ is white noise. Equation (8) becomes for this case

$$\begin{aligned}\frac{d\theta_{00}}{dt}-2\theta_{10}&=0,\\ ab\theta_{00}+\frac{d\theta_{10}}{dt}+(a-b)\theta_{10}-\theta_{11}&=0,\\ 2ab\theta_{10}+\frac{d\theta_{11}}{dt}+2(a+b)\theta_{11}&=2\phi_0 b,\end{aligned} \tag{12}$$

which must be solved subject to the initial conditions $\theta_{00}=0$, $\theta_{10}=0$, $\theta_{11}=\phi_0$. The solution for $\langle x^2(t)\rangle_{Av}$ in this case is

$$\langle x^2(t)\rangle_{Av}=\theta_{00}=\phi_0\left[\frac{1}{a(a+b)}+\frac{e^{-2at}}{a(a-b)}-\frac{2e^{-(a+b)t}}{a^2-b^2}\right]. \tag{13}$$

ACKNOWLEDGMENTS

The author would like to acknowledge the contributions made by Dr. N. F. Parker and Dr. R. E. Roberson in discussions of this problem.

JOURNAL OF APPLIED PHYSICS VOLUME 24, NUMBER 5 MAY, 1953

Temperature Dependence of Low Frequency Fluctuations in Thermionic Emitters

T. B. TOMLINSON
Department of Electronics, University of Southampton, England
(Received November 21, 1952)

A method is described for measuring the temperature variation of the low frequency fluctuations of emission current (flicker effect) in diodes with oxide-coated cathodes. In order to vary the cathode temperature and at the same time avoid the effects of space charge and the variation due to changes of the mean current, it is necessary to operate the test diodes under retarding field conditions. The results show that the majority of tubes exhibit minimum fluctuations at a cathode temperature in the region of 1000°K, although both rising and falling temperature characteristics are encountered. It is concluded that the flicker effect is in fact a combination of at least two separate phenomena. Finally, the various theories are examined in the light of the new information, and none is found to be completely satisfactory.

INTRODUCTION

IN a previous report[1] experiments were described which proved that the low frequency fluctuations of current in thermionic emitters are reduced by space charge to a greater extent than are the normal shot noise fluctuations. In these experiments, it was necessary to operate the test diodes under retarding field conditions, because the magnitude of the low frequency fluctuations (flicker effect) was found to be dependent on the temperature of the cathode. At the conclusion of the above work, it was decided to examine this temperature dependence in greater detail in order to obtain new information on the origin of flicker noise. The following is an account of this investigation and the results obtained.

OUTLINE OF METHOD

It is well established that the mean square value of the low frequency noise current for a given band width increases both with increasing mean current and with decreasing midfrequency of the pass band. Consequently, all measurements must be made at a given current and at the same midfrequency. Furthermore, the noise current to be measured must not be reduced by space charge which would obscure the variation due to temperature. These conditions can be satisfied over a range of cathode temperatures only by operating the diode under retarding field conditions.[1] An unavoidable disadvantage is the restriction of the measurements to low current levels; this is especially true at lower cathode temperatures where only a very small current range is available under true retarding field conditions. As the flicker noise component of the fluctuations is dependent on the mean current to a higher degree than the shot noise component, the latter may be the larger of the two unless a sufficiently low midfrequency is chosen. On the other hand, a lower midfrequency leads to a narrower band width, and this increases the recording time

[1] T. B. Tomlinson, J. Appl. Phys. 23, 894 (1952).

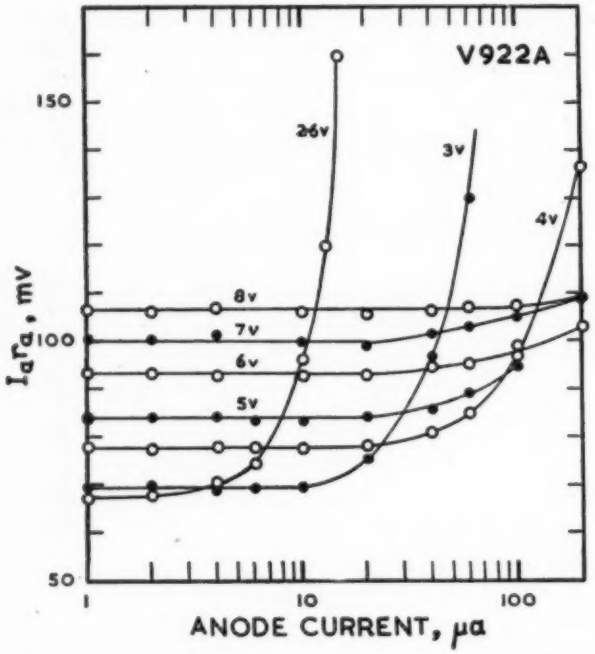


FIG. 1. Characteristics of a test diode when operated in the retarding field region. For low currents the product $I_a r_a$ has a constant value which is a measure of the cathode temperature.

necessary to achieve a given accuracy of measurement of the mean square value. In addition, a limit is set to the sensitivity and accuracy by the masking effect of the noise introduced by the following amplifier tubes and circuit elements. As a compromise, a midfrequency of 20 cps was selected and the test diodes operated at a mean current of 2 μa. For the average tube this gave a temperature range of some 800°K to 1300°K in which retarding field conditions could be maintained. The upper temperature limit is set by the rapid rate of evaporation of the oxide layer from the emitting surface and the danger of burning out the heater filament.

In order to calculate the fluctuations of anode current (zero load) from the fluctuations of voltage measured at the anode, the diode internal resistance r_a had to be evaluated for the selected working conditions. The ac bridge method was again used for this purpose.

MEASUREMENT OF CATHODE TEMPERATURE

The temperature of a cathode is not uniform over the whole surface, therefore, it will be realized that any measurement of temperature is relative rather than absolute. Also, no special diodes were available for these tests, so that use had to be made of the best of the known indirect methods.

The bridge measurements of r_a were extended to higher values of anode current I_a and the product $I_a r_a$ plotted against corresponding values of I_a (Fig. 1). For a given cathode temperature, the product $I_a r_a$ is constant over a limited range of currents, where the diode is working under true retarding field conditions and increases as soon as space charge limitation of current sets in. The value of $I_a r_a$ over the region in which it is constant is a measure of the temperature of the cathode.[2]

$$kT = e \cdot I_a r_a$$

where e = electronic charge and k = Boltzmann's constant. It was particularly convenient to use the above method, because the apparatus for measuring r_a was already available.

For confirmation, readings were taken of the heater supply voltage and current so that the heater power P_h corresponding to the various measured r_a values could be calculated. According to Widdell[3] the cathode temperature is proportional to $P_h^{\frac{1}{4}}$; by plotting $I_a r_a$ against $P_h^{\frac{1}{4}}$ the experimental points were found to lie closely distributed about a straight line through the origin (Fig. 2). As the accuracy of the $P_h^{\frac{1}{4}}$ values was within ½ percent, corrected values of $I_a r_a$ and the corresponding temperatures were deduced from the straight line relation. In view of the likely variation along the cathode, the figure for the temperature obtained in this manner is probably more realistic from the emission point of view than any derived from a direct method.

MEASUREMENT OF FLUCTUATIONS

The test vacuum tube was a normal type indirectly heated diode with oxide-coated cathode; the only special requirement was that the physical dimensions of the anode and cathode would be suitable for the maintenance of retarding field conditions down to low temperatures.[2] A selection of close-spaced diodes was made with this requirement in view.

The apparatus used for measuring the fluctuations of the diode current was the same as that described previously,[1] so only a brief description will be included here. The noise voltage appearing across the diode load resistor was amplified by a high gain selective amplifier having a stable over-all frequency band width of 1.7 cps

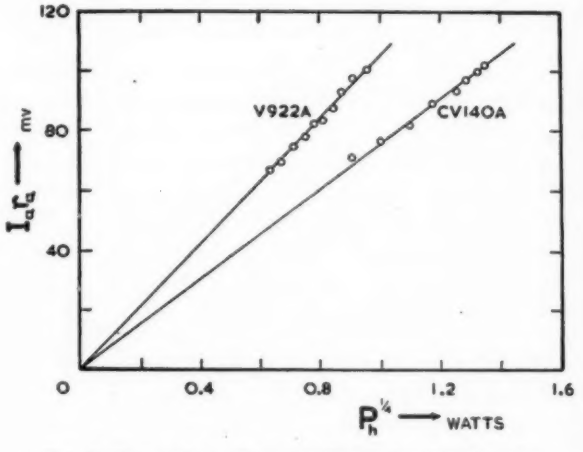


FIG. 2. Comparison of two methods of estimating the cathode temperature.

[2] R. Fürth and D. K. C. MacDonald, Proc. Phys. Soc. (London) 59, 375 (1947).
[3] E. G. Widdell, RCA Rev. 5, 106 (1940).

centered on 20 cps. The mean square value of the amplified noise voltage was measured by means of a thermojunction whose output was fed via a dc amplifier to a recording milliameter. The signal level at the junction was kept sufficiently low for a square law characteristic to apply. The thermojunction and subsequent stages were calibrated as a whole for ac signals, and the mean square value was obtained by visually averaging the noise recording.

The filament of the test diode was fed via a potentiometer from two large capacity 6v storage cells which, after an initial discharge period, could be relied upon to give a steady filament voltage over a long time interval. The circuit was connected so that the mean dc potential of the heater was positive with respect to the cathode. This insured that no direct emission from heater to anode could occur; the error due to additional noise introduced by such emission could be appreciable at the low anode current concerned.

Calibration

To obtain an absolute measure of the noise output voltage, its magnitude was compared with that of the thermal noise voltage appearing across resistors of known value connected across the input terminals of the first amplifier stage. A recording was also taken with these terminals shorted. As a refinement on the previous method the mean square output voltage V^2, as measured by the recorder, was plotted against the total resistance R_0 in the input circuit including the parallel grid leak of the first amplifier tube. According to the Nyquist formula,

$$V^2 = 4kTR\int_0^\infty G_f^2 \cdot df$$

in which T is the temperature (°K) and G_f is the overall voltage gain of the amplifying system expressed as a function of frequency f. The mean square output voltage V^2 was plotted against resistance R_0 and a straight line graph was obtained (Fig. 3). By extrapolating the straight line back to zero noise level the equivalent noise resistance R_{eq} of the first amplifier and subsequent stages was read off directly. Its value was found to be 65kΩ. This method of calibration avoids the evaluation of the integral $\int_0^\infty G_f^2 \cdot df$ as will be seen later.

Experimental Procedure

The heater supply voltage to the test diode was adjusted to give the required cathode temperature, and the anode supply voltage was adjusted to give the required 2 μa anode current. After a settling down period of 20–40 minutes, the noise output voltage V_v^2 was recorded over an interval of some 20 minutes. Recordings were made at progressively higher temperatures up to the maximum. The temperature was then reduced to the original values and recordings taken again in order to see if repeatable results had been obtained. It appeared that the results were affected by the previous operating temperature of the cathode especially in the low temperature range. The effect was most noticeable in the V922A, but in no case was it sufficiently pronounced to obscure the general trend of the temperature variation curve.

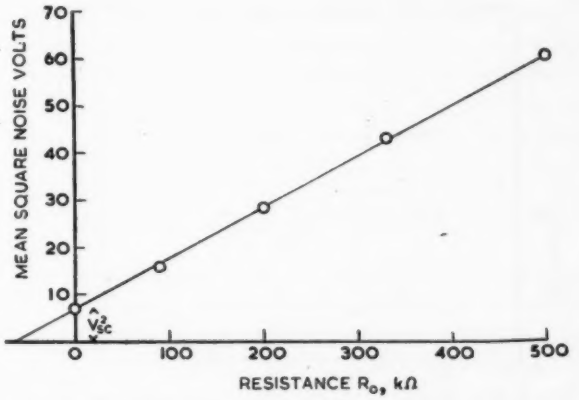


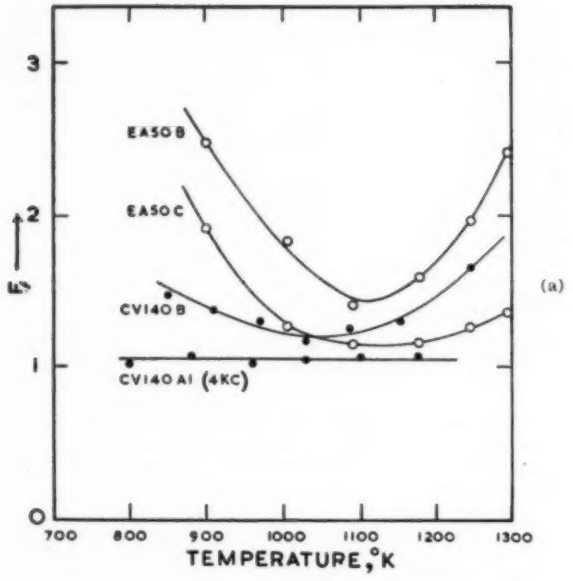
FIG. 3. Calibration curve for the amplifying and recording equipment. By extrapolation back to zero noise volts, the equivalent noise resistance R_{eq} of the amplifier and associated circuits is found to be 65kΩ.

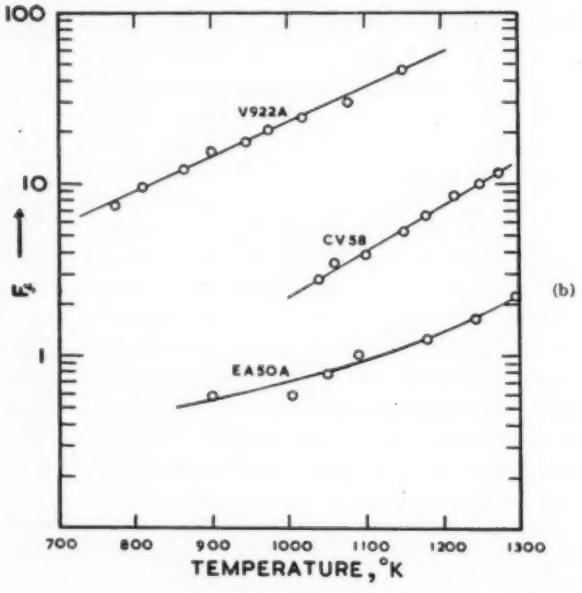
An attempt was made to reduce the discrepancy by allowing the cathode to cool down between successive measurements, but this was not successful. Apparently the state of the cathode coating is "frozen in" during inactive periods. Consistent results were obtained by setting up a separate chassis in which the test valves were caused to draw the 2-μa current at the selected cathode temperature for a period of several hours before being transferred to the test equipment in which they were then operated under identical conditions. By this procedure, the temperature plots of several diodes were taken concurrently, and this greatly speeded up the process.

As a precautionary measure all diodes were tested for any additional noise due to small leakage currents across the micas or glass envelope. For this purpose, a noise recording was taken with the cathode hot but with the anode voltage sufficiently negative to prevent the flow of anode current. The diode was then removed from the test apparatus, and the new noise voltage recorded. Unless the new mean square value was closely equal to that of the first recording, the specimen diode was discarded. Only one such defective diode was discovered; it was noticed that the additional noise due to leakage produced a less regular, more peaky type of recording.

Calculation

The mean square noise output voltage V_{sc}^2 contributed by the amplifier was read off from Fig. 3. Then $V_v^2 - V_{sc}^2$ is the total mean square noise voltage originating across R, the effective diode load. R consists of the parallel combination of the diode load resistor and





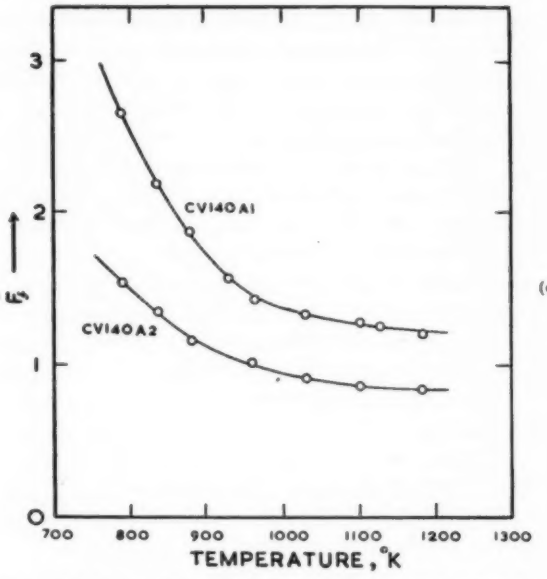


FIG. 4. (a), (b), (c). Observed variation of the flicker noise factor F_f with cathode temperature for a number of test diodes. Curves (a) are the most usual type encountered.

the amplifier grid leak and had a value of 90kΩ. A correction must be made for the thermal noise V_R arising in R. By direct proportion $V_R=V_{sc}\cdot R/R_{eq} \simeq 1.38V_{sc}$. Due to the shunting effect of the diode, internal resistance the thermal noise voltage set up across the anode load circuit is reduced by a factor $r_a/(r_a+R)$, so that the mean square output voltage solely due to fluctuation of anode current is

$$V_v^2-V_{sc}^2-1.9V_{sc}^2r_a^2/(r_a+R)^2.$$

Therefore, the mean square value $\langle\delta I_a\rangle_{kf}^2$ of the anode current fluctuations (zero load) per unit band width is

$$\frac{V_v^2-V_{sc}^2-1.9V_{sc}^2r_a^2/(r_a+R)^2}{\int_0^\infty G_f^2\cdot df\cdot r_a^2R^2/(r_a+R)^2}.$$

From the calibration,

$$V_{sc}^2=4kTR_{eq}\int_0^\infty G_f^2\cdot df$$

so that

$$\langle\delta I_a\rangle_{kf}^2=\left\{(V_v^2-V_{sc}^2)\left(\frac{r_a+R}{r_aR}\right)^2-\frac{1.9V_{sc}^2}{R^2}\right\}\cdot\frac{4kTR_{eq}}{V_{sc}^2}.$$

It is convenient to express the anode current fluctuations in terms of the equivalent shot noise fluctuations $\langle\delta I_s\rangle_{kf}^2$ for the same anode current, i.e., $\langle\delta I_a\rangle_{kf}^2=F^2\langle\delta I_s\rangle_{kf}^2$, where F is a factor to be evaluated. For unit band width,

$$\langle\delta I_s\rangle_{kf}^2=2eI_a$$

so that

$$F^2=\left\{\left(\frac{V_v^2}{V_{sc}^2}-1\right)\left(\frac{r_a+R}{r_aR}\right)^2-\frac{1.9}{R^2}\right\}\cdot\frac{2kTR_{eq}}{eI_a}.$$

The term $2kTR_{eq}/eI_a$ is a constant and need be evaluated once only, although the calibration must be tested from time to time. The shot noise component included in $\langle\delta I_a\rangle_{kf}^2$ has yet to be taken into account. If the mean square value of the current fluctuations arising solely as flicker effect is $\langle\delta I_f\rangle_{kf}^2$ and the corresponding factor is F_f^2, then

$$\langle\delta I_f\rangle_{kf}^2=\langle\delta I_a\rangle_{kf}^2-\langle\delta I_s\rangle_{kf}^2$$

and

$$F_f=(F^2-1)^{\frac{1}{2}}.$$

RESULTS

A total number of 17 diodes were investigated in the manner described and the measured values of F_f were

plotted against the corresponding cathode temperatures. (Figure 4.) The curves (a) are examples of the variation most commonly encountered; they show F_f to have a minimum in the region 1000°K to 1100°K and to increase by some 100 percent at the lowest and highest temperatures observed. The curves (b) indicate an exponential increase of F_f with increase of temperature; these are the only results of this type that were observed. In comparison the curves (c) show a preliminary fall of F_f with increasing temperature and a tendency to level out near 1200°K.

As a test of the method, a set of results was obtained at the higher frequency of 4 kcps for the diode CV140 Al. At this frequency and at the 2 μa anode current the flicker effect should be negligible compared with the shot effect. The latter phenomenon concerns fluctuations of purely statistical origin, so that no temperature variation should be apparent. The results obtained are graphed in Fig. 4(a) and show F to be close to the expected value of unity at all temperatures.

CONCLUSIONS

The results suggest that two separate effects are involved in low frequency fluctuations of emission from oxide coated cathodes. One of these has a rising temperature characteristic and is the dominant effect in the curves (b), the other has a falling temperature characteristic and is dominant in the results (c). In the most common case, as exemplified by the results (a), the two effects are of similar order and a characteristic showing a minimum value of F_f is observed.

The various theories of the flicker effect may now be examined more critically in the light of these results. In the retarding field condition there is no potential minimum due to the space charge and the point of lowest potential is at the anode. The anode current is then dependent only on the initial velocities of the emitted electrons, the anode-cathode potential difference and the work function of the anode material. It cannot be changed by the alteration of the potential barrier at the emitting surface unless such alteration causes the height of this barrier to rise above that at the anode surface. In Macfarlane's earlier theory[4] the fluctuations of current were assumed to be due to surface changes including the formation at the surface of a dipolar layer which lowered the potential barrier in that region. It is clear that this model would not account for fluctuations under retarding field conditions.

In a second theory of contact noise, Macfarlane[5] introduces an increased potential barrier at the surface due to the presence of negative ions. As stated above, this effect could be an explanation of the presence of flicker noise under retarding field conditions provided that only increases of the surface barrier above the level of the anode potential barrier were considered. A detailed examination of the modified theory is in preparation.

In Schottky's theory,[6] the fluctuations were assumed to be due to the influence of foreign atoms in the surface in accordance with Langmuir's doublet theory. The emission over a small surface region could either be increased or decreased according to the surface condition. Again a modification to the theory is necessary in order to include the retarding field conditions. The changes in the effective work function of the cathode must be restricted to increases only and these must be of such magnitude as to momentarily cause the potential barrier at the cathode to be the determining factor.

Van der Ziel[7] has suggested that more than one form of flicker effect is likely to be present in oxide-coated emitters, and this is in agreement with the above experimental observations. He describes a new model in which fluctuations of current are the result of spontaneous fluctuations in the conductivity of the semiconducting material which forms the activated oxide layer. In one respect this theory is incorrect; it is assumed that the potential difference across the oxide coating does not fluctuate. However, qualitative tests using a special diode with a probe embedded in the oxide coating show that a noise voltage is set up between the probe and the cathode, even when the mean current through the probe contact is zero. The mean square value of this noise voltage is independent of the mean anode current flowing through the coating but increases as the temperature is reduced. It would seem certain that this fluctuation of potential difference across the oxide coating must influence the anode current so that a noise component which decreases with increase of the cathode temperature is to be expected. This mechanism could explain the results of Fig. 4(c).

One concludes that no completely satisfactory explanation for the flicker effect has so far been given and that further experimental investigation of the spontaneous variations of conductivity of the oxide coating may yield useful information on the origin of the current fluctuations.

[4] G. G. Macfarlane, Proc. Phys. Soc. (London) **59**, 366 (1947).

[5] G. G. Macfarlane, Proc. Phys. Soc. (London) **B63**, 807 (1950).

[6] W. Schottky, Phys. Rev. **28**, 74 (1926).

[7] A. Van der Ziel, Physica **16**, 359 (1950).

The X-Ray Shadow Microscope

V. E. Cosslett and W. C. Nixon*
Cavendish Laboratory, Cambridge, England
(Received December 30, 1952)†

Shadow projection permits higher resolution and simpler operation than other methods of x-ray microscopy. A point source, less than 1μ in diameter, is obtained by the use of magnetic electron lenses and used for projecting shadow images at a primary magnification up to 100× with an exposure time in the order of minutes. The use of a metal foil as target and tube window insures high intensity at the specimen, which is in air. A resolving power little short of the best optical performance has been obtained with an accelerating voltage of 10 kv. The practical limitations to resolution, contrast, and exposure time are discussed in detail. A resolution of 100A should be practicable, but only with longer x-rays and an exposure of the order of 1 hour.

1. INTRODUCTION

A SUITABLY designed microscope, using x-rays, would give images of adequate contrast at a resolution much better than the 1000A of the best optical microscope, and without the specimen having to be placed in a vacuum as in electron microscopy.

Other advantages of x-ray microscopy may be listed as (a) greater penetration of the specimen than is possible with light or with electrons; (b) determination of minute masses in the specimen by the known laws of x-ray absorption;[1] (c) use of heavy metal stains to accentuate contrast in selected parts of the specimen.

In addition, with a shadow type of microscope, all the specimen is in focus at once and stereographic presentation of the image is possible, even at the highest magnification.

The simplest method of obtaining x-ray micrographs is to place the specimen in contact with a photographic film of the maximum resolution type and expose it to the radiation from a normal x-ray tube. All the enlargement is obtained photographically and the resolution is primarily limited by the grain size of the emulsion.

Several workers[1,2] have obtained micrographs with a resolution of 1μ, but in any case the resolution cannot exceed that of the optical system used to enlarge the x-ray negative, even if a grainless emulsion were available.

A second and more difficult method is to use curved crystals as x-ray mirrors and construct an x-ray microscope similar to the reflecting optical microscope used for ultraviolet light. An original theoretical limit of resolution of 70A claimed[3] for this method has since been revised to 2000A (0.2μ).[4,5] A resolution of 1μ has been reached experimentally by a few instruments of this type.[6]

The third possibility is to make an x-ray shadow microscope, as first proposed by Sievert.[7] A practical system has been constructed for the first time in the present work, giving a resolution less than 0.5μ with an exposure of 5 min. In this method, a thin metal foil 1μ thick, which forms part of the vacuum wall of the x-ray tube, also acts as the source of x-rays.[8] Magnetic electron lenses focus an intense beam of electrons on to a small area of this target, producing a point source of x-rays which casts an enlarged shadow image of a nearby object on to a distant photographic plate. Electron probes of 100–200A have been produced for other purposes,[9] so that the theoretical limit of resolution of this type of x-ray microscope approaches that of the electron microscope, while the specimen is still at atmospheric pressure.

A brief initial account of this shadow method has recently been amplified by a description of some applications in biology, especially in entomology.[10] The present contribution is concerned with the physical aspects of the method, the construction and operation of the apparatus, and a detailed discussion of its limitations.

2. PHYSICAL PRINCIPLES

The arrangement of the apparatus is shown in Figs. 3–5. Electron lenses are employed to form a greatly reduced image of a tungsten filament on a metal foil which seals the evacuated tube. The fine x-ray focus thus formed projects a shadowgraph of the specimen on to a recording emulsion at a distance. Limitations arise of two kinds, which need separate discussion. First, the size of the electron spot is limited by the finite size of the source and by the aberrations of the final electron lens; in turn, the spot size limits the resolution in the x-ray shadowgraph. Second, the intensity of the emitted beam of x-rays is limited by the emissive power of the electron source and by the rate of energy dissipa-

* Present address: General Engineering Laboratory of the General Electric Company, Schenectady, New York.
† The original manuscript was received August 11, 1952.
[1] A. Engstrom, *Progress in Biophysics* (Butterworth Company, London, 1950), p. 164.
[2] W. Ehrenberg and W. E. Spear, Nature **168**, 513 (1951).
[3] P. Kirkpatrick and A. V. Baez, J. Opt. Soc. Am. **38**, 766 (1948).
[4] W. Ehrenberg, J. Opt. Soc. Am. **39**, 741 (1949).
[5] E. Prince, J. Appl. Phys. **21**, 698 (1950).
[6] C. M. Lucht and D. Harker, Rev. Sci. Instr. **22**, 392 (1951).
[7] R. Sievert, Acta Radiologica **17**, 299 (1936).
[8] M. von Ardenne, Naturwiss. **27**, 485 (1939); V. E. Cosslett, Research Project, London University (1939).
[9] J. Hillier and R. F. Baker, J. Appl. Phys. **15**, 663 (1944).
[10] V. E. Cosslett and W. C. Nixon, Nature **168**, 24 (1951); V. E. Cosslett and W. C. Nixon, Proc. Roy. Soc. (London) **B140**, 422 (1952).

tion in the target foil. Other quantities, such as target thickness and accelerating voltage, are fixed within narrow limits by practical considerations. The recording conditions (contrast and exposure time) depend on the two primary limiting factors, and will be discussed subsequently.

Resolving Power and Size of Source

The size of the electron spot on the target governs the size of the x-ray source and hence the resolution attainable. The shadow image will have an unsharpness caused partly by the penumbra resulting from the finite size of the source of x-rays, and partly by Fresnel diffraction, as shown in Figs. 1 and 2 for the simple case of an opaque straight edge E illuminated from a source S of finite diameter s, at a distance b from it forming a shadow image on a screen A at a distance a.

The spacing p of the Fresnel diffraction fringes formed on the screen (Fig. 1) is given by the optical relation

$$p=a/b\cdot\nu\cdot(\lambda b(a-b)/2a)^{\frac{1}{2}},$$

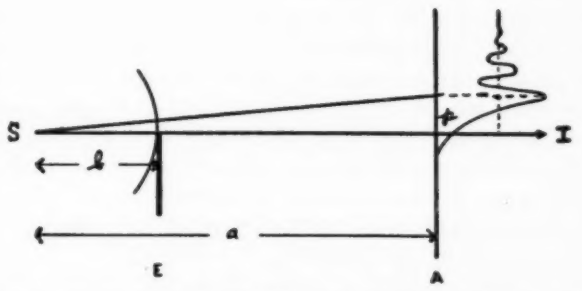


FIG. 1. Formation of Fresnel fringes at straightedge E illuminated by point source S.

where λ is the wavelength. The value of ν is given by the Fresnel integrals, and is approximately $\sqrt{2}$ for the first intensity maximum, measured from the geometrical shadow of the edge. Taking this distance as equivalent to the resolution at the screen, the least resolved distance in the object becomes

$$\delta_F=p/M=(\lambda b)^{\frac{1}{2}}, \qquad (1)$$

on writing $(a-b)\doteq a$. As the wavelength is fixed by contrast considerations, the object distance must be kept as small as possible. For $\lambda=2.4A$ and a resolved distance of 1μ we find $b\leq 4$ mm, but for $\delta_F=0.1\mu$, $b\leq 40\mu$. A resolving power of optical order thus requires precise location of the specimen with respect to the focal spot; this is practicable when the spot is formed on the window of the x-ray tube.

If Fresnel diffraction is made negligible in this way then the finite size of the x-ray source becomes the limiting factor. A penumbra of width p' extends from the edge of the umbra (Fig. 2), and two neighboring edges will no longer be distinguishable if their respective umbras fall closer together than p'. In the object space

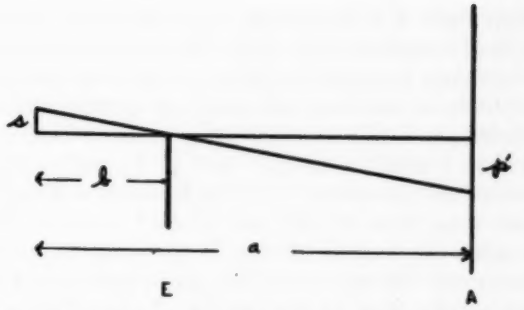


FIG. 2. Effect of finite size (s) of source in blurring image of straightedge E received on screen A.

we have $\delta_P=p'/M=s(a-b)/a\doteq s$, since $a\gg b$. Hence, the least resolved distance in the x-ray image cannot be smaller than the spot size.

The size of the x-ray source may be taken as that of the exciting electron spot, if the latter is larger than the range of electrons of given voltage in the target foil and if the foil is thin, compared with the range of the generated x-rays. The target metal must be of high atomic number, to give high efficiency of x-ray production, and the voltage must be low, to give x-rays sufficiently soft to be absorbed in biological material. An operating voltage of 10 kv gives an electron penetration in tungsten of 0.2μ; i.e., at this distance the most probable terminal velocity is zero.[11] The x-rays produced have a minimum wavelength of 1.2A and a maximum intensity at a wavelength between 2 and 4A. Such radiation is negligibly absorbed in tungsten foil of thickness similar to the electron penetration. So long as the size of the electron spot is larger than 0.2μ, the x-ray spot can be assumed to be of the same size. The smallest spot so far obtained is about twice this value.

The electron spot is limited in size by the geometry of the electron optical system and by the aberrations of the focusing fields. In first approximation the image radius r_2 is given by the thin lens expression $r_2=r_1\cdot v/u$, where r_1 is the radius of the effective source, u, its distance from the lens and v, the image distance. The effective value of r_1 will be the radius of the emitting cathode as seen from the lens, which will differ from that of the actual emitting area owing to aberrations in the electric fields in front of it. Suitable operating conditions can be chosen[12] to minimize these effects; very fine wire is used for the cathode filament, to make r_1 small ($<50\mu$). The image size then depends on the ratio v/u, made small by placing the filament a long way from the lens and using a lens of shortest possible focal length, so that $v\doteq f$. Two considerations set a limit to the demagnification obtainable with a single lens. Since the permissible aperture is fixed by spherical aberration the gathered intensity falls off in proportion to u^2, and exposure times become

[11] R. Whiddington, Proc. Roy. Soc. (London) **A89**, 554 (1914); H. M. Terrill, *X-Ray Technology* (D. Van Nostrand Company, Inc., New York, 1930).

[12] M. E. Haine and P. A. Einstein, Brit. J. Appl. Phys. **3**, 40 (1952).

unduly great if u is made too large. Secondly, the focal length of a magnetic lens cannot be indefinitely reduced, owing partly to saturation effects in the iron pole pieces and partly to practical restrictions on the minimum size of their bore. It becomes difficult to obtain a focal length less than 1 mm.[13] A focused spot of $\frac{1}{2}\mu$ radius can be obtained from a source of radius 50μ with a demagnification ratio $1/m$ of 100. For $v \doteqdot f = 1$ mm the object distance must then be 100 mm; in the present apparatus it was made 380 mm to permit the formation of a spot much smaller than 1μ in diameter. The condenser lens between source and objective lens allows the effective value of u to be varied.

Spherical Aberration

The actual size of the cathode image will be larger than calculated for paraxial conditions, owing to aberrations in the objective lens. The relative importance of these has been well established by electron microscopy. The chromatic error may be easily reduced below the resolution here sought by providing well smoothed electrical supplies; a stabilization of 1 part in 1000 gives a chromatic disk of confusion less than 0.1μ in diameter in normal operating conditions. Astigmatism is reduced by insuring the highest precision in machining the lens bore; an accuracy of 5μ is required to reduce the diameter of the astigmatic disk below 1μ.[14] Any residual astigmatism caused by errors in machining or by inhomogeneity of the iron could be compensated with radial rods.[15] Space charge effects are too small to produce measurable spreading of the beam, even at the high current density obtained in a spot of 0.1μ diameter.

The chief restriction is imposed by spherical aberration, as no means are known of correcting it in electron lenses. It produces a disk of confusion of radius δ_s given by

$$\delta_s = C_s \alpha_2^3, \tag{2}$$

where α_2 is the semi-aperture of the beam in the image space and C_s is a factor linear in dimensions and of the order of the focal length. Hence a physical aperture must be fitted in the lens to reduce α_2 below the value for which $\delta_s = r_2$. For $C_s = f = 1$ mm and $r_2 = 0.5\mu$ the maximum permissible aperture $= 8 \cdot 10^{-2}$ radian, i.e., a physical aperture of radius 80μ at the principal plane of the image space. The smallest stop used in the present work has been of radius 35μ, equivalent to $2\delta_s = 2r_2 = 0.1\mu$; with $C_s = 1$ mm, Eq. (2) then gives $\alpha_2 = 3.7 \cdot 10^{-2}$.

Energy Balance in the Target

In x-ray tubes of conventional design, the attainment of maximum output intensity requires forced cooling of the target since the rate of heat dissipation is less than the rate of energy supply from the electron beam. However, it was shown by Goldsztaub[16] and by Oosterkamp[17] that as the spot size is reduced the conditions of energy equilibrium change in such a way that greater specific loading becomes permissible. With a very small spot the rate of dissipation of heat is relatively so high that the limitation on x-ray intensity becomes the specific emission of the electron source.[18]

For a target cooled on the side far from the electron spot and for a spot large compared with the thickness of the target (t), but small compared with its radius, heat flow is almost entirely along the normal from the front to the rear surface (axial flow). The permissible specific loading W_x is

$$W_x = k(T_1 - T_2)/t, \tag{3}$$

where k is the thermal conductivity, T_2 the temperature of the cooled surface and T_1 the maximum permissible temperature in the spot, usually taken as 0.75 of the melting temperature of the target metal. When the spot radius r_2 is of the same order as t, radial heat flow becomes appreciable and may finally be more important than axial flow. Oosterkamp[17] has calculated a correction factor for Eq. (3), to take account of this change. For a spot of diameter 1μ on a copper target 1μ thick, W_x is then 680 kw/mm^2, compared with specific loads of the order of 1 kw/mm^2 in the best designs of x-ray tubes with spots of normal size. The energy flux, $E = \pi r_2^2 \cdot W$, proves to be approximately proportional to r_2, so that total x-ray output falls proportionally as the spot size is reduced. For a 1μ spot and 1μ target thickness, $E_x = 0.50$ watt on copper; on tungsten, $E_x = 0.60$ watt.

The rate of supply of energy E_s is determined by the emissive power of the electron source and by the gathering aperture of the objective lens. The former is measured by the product of the current emitted per unit area per unit solid angle (β) and the accelerating voltage V_2. As in an optical system, the brightness β is invariant in the absence of absorption or scattering. Also the total current is constant, $j_1 r_1^2 = j_2 r_2^2$, if j_1 and j_2 are the current densities at object and image, respectively. Expressing the optical invariance rule in the Helmholtz-Lagrange form: $r_1 n_1 \sin\alpha_1 = r_2 n_2 \sin\alpha_2$, where n_1 and n_2 are the refractive indices of object and image space, determined by the respective potentials V_1 and V_2, then

$$j_2/j_1 = r_1^2/r_2^2 = n_2^2 \sin\alpha_2 / n_1^2 \sin\alpha_1 = \frac{V_2}{V_1}\alpha_2^2, \tag{4}$$

since $\alpha_1 = \pi/2$ and $\sin\alpha_2 \doteqdot \alpha_2$ in the experimental conditions. V_1 is given by the velocity of thermal emission of electrons (about $\frac{1}{4}$ V at the filament temperatures here used), and V_2 is the accelerating voltage between cathode and anode, usually 10 kv. The remaining quantity α_2 is fixed by the necessity of minimizing the

[13] G. Liebmann and E. M. Grad, Proc. Phys. Soc. (London) **B64**, 956 (1951).
[14] P. A. Sturrock, Trans. Roy. Soc. (London) **A243**, 387 (1951).
[15] J. Hillier and E. G. Ramberg, J. Appl. Phys. **18**, 48 (1947).
[16] S. Goldsztaub, Compt. rend. Acad. Sci. (Paris) **224**, 458 (1947).
[17] W. J. Oosterkamp, Philips Tech. Rev. **3**, 303 (1948).
[18] V. E. Cosslett, Proc. Phys. Soc. (London) **B65**, 782 (1952).

effect of spherical aberration (Eq. (2)). For a resolution of 1μ, $\alpha_2=8\cdot10^{-2}$ for a lens of $f=1$ mm, so that Eq. (4) yields $j_2=256j_1$.

The current density incident at the target can now be related to the rate of energy dissipation (Eq. (3)). Since $E_x=0.60$ watt for a spot of 1μ diameter, the permissible maximum current density at a tungsten target for $V_2=10$ kv, will be $j_2=7500$ amp/cm². This would call for a specific emission from the source of $j_1=30$ amp/cm², which is much higher than that given by a filament at ordinary operating temperature (~1 amp/cm²). It might be obtained by raising the temperature, but only at the cost of reducing the life of the filament to a value smaller than the exposure time needed for a single micrograph. In practice the beam intensity is less than the value which would raise the target to the maximum permissible temperature, that is, the specific emission of the electron gun replaces the energy balance in the target as the limiting factor in operating the x-ray tube. With the objective lens operating at maximum power, further increase in resolution requires reduction in α_2 to satisfy $\delta_s=r_2$, and the amount of energy that can be delivered to the target diminishes rapidly while the rate at which it can be dissipated increases. For a resolution of 0.1μ, the target could dissipate a maximum intensity $j_2=7.5\cdot10^4$ amp/cm². But a filament operated at 10 amp/cm² could only supply, through the required aperture $\alpha_2=3.7\cdot10^{-2}$, an intensity $j_2=55j_1=550$ amp/cm².

Efficiency of X-Ray Production and Absorption

The proportion of the energy in the electron beam that is converted into x-rays depends on the accelerating voltage V_2 and the atomic number Z of the target material. To a close approximation the efficiency ϵ is given[19] by

$$\epsilon=1.1\cdot10^{-9}\cdot ZV_2=1.35\cdot10^{-6}Z/\lambda_m, \qquad (5)$$

where λ_m is the minimum wavelength of the generated x-rays (in A). It is thus desirable to work with as high a voltage as possible, but considerations of contrast require the converse. The absorption coefficient τ for soft x-rays in biological material is approximately given[13] by

$$\tau=k\cdot Z^4\cdot\lambda^3, \qquad (6)$$

so that the wavelength should be as long as possible, the voltage as low as possible. As a compromise between efficiency and contrast, a voltage was selected which gave an intensity sufficient for a reasonably short exposure and a wavelength long enough for adequate contrast in biological specimens. Experience showed that a reasonable degree of contrast was obtained with $\lambda\sim2$A, requiring a voltage (12 kv) corresponding to $\lambda_m\sim1$A because of the shape of the intensity-wavelength distribution.

To minimize loss by absorption of such short x-rays in the target, it must not be thicker than is required for efficient absorption of the electron beam and for mechanical strength to withstand atmospheric pressure. A tungsten foil 1μ thick‡ is strong enough if supported on a thick washer with a very small exposed area; an opening of diameter 250μ (0.010 inch) was used in most of this work. The absorption loss of 2A x-rays in 1μ of tungsten is 44 percent. When a spot smaller than 1μ diameter is employed, with correspondingly weaker x-ray output, it is clearly desirable to use a thinner foil.

The camera must similarly be kept short to minimize absorption of x-rays in the atmosphere. With object distance $b=1$ mm and $M=40$, the camera length $a=40$ mm and the absorption loss is 10 percent; a has usually been in the range 30–80 mm.

Considerations of efficiency also govern the choice of target metal, since maximum production of x-rays is required (Eq. (5)) and maximum dissipation of heat (Eq. (3)). On the first ground a metal of high Z is indicated, and on the second one of high thermal conductivity. A survey of the data[20] shows that gold has the required combination of properties in the highest degree, followed by silver, with copper and tungsten equal third. When the melting point is taken into account also, the order of preference becomes tungsten, gold, silver, copper. In the present work tungsten foil has usually been used as target.

Recording the Image

The x-ray shadowgraph is recorded on a photographic emulsion at a suitable distance from the source. The choice of distance is conditioned by several considerations: diffraction effects, exposure time, the need for a large primary magnification, and photographic grain size. The last is the decisive factor as the resolving power is not usefully employed unless the primary magnification is at least enough to ensure that two separated points in the image do not fall on the same grain; if the plate is placed too close to the object, resolved detail will not be distinguished in the negative. Since the grain size in the usual type of x-ray emulsion is of the order of 20μ, a minimum primary magnification M of 20 times is needed at a resolution of 1μ; a further photographic enlargement of 10× will make the resolved points directly visible, assuming the resolving power of the unaided eye to be 200μ. As the grain of the emulsion might then be obtrusive, it is desirable to arrange a higher primary magnification, say 40×, with a correspondingly smaller final enlargement. The maximum object distance b for a resolution of 1μ is 4 mm (Eq. (1)), but this should be reduced in the interest of short exposure. With $b=1$ mm, the camera length would be 40 mm, small enough to give an exposure time of the order of 1 min with maximum target loading on a spot of

[19] A. H. Compton and S. K. Allison, *X-Rays in Theory and Experiment* (D. Van Nostrand Company, Inc., New York, 1935), p. 90.

‡ Provided by Philips Research Laboratories, Eindhoven, Holland, through the kindness of W. J. Oosterkamp.

[20] W. J. Oosterkamp, Philips Tech. Rev. 3, 58 (1948).

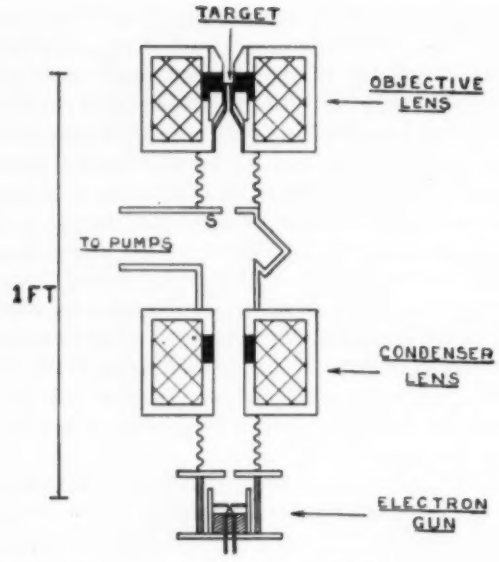


FIG. 3. Schematic diagram of x-ray shadow microscope.

diameter 1μ. For higher resolution, b has to be reduced proportionally to r_2^2 (Eq. (1)), while the primary magnification (a/b) should be varied in inverse proportion to r_2. A resolution of 0.1μ requires $b \leq 40\mu$ and $M \geq 200$, which demands $a \doteqdot 8$ mm, a length which begins to be inconveniently small.

The exposure time is not at first radically altered as the size of spot is reduced: As the power of the objective is raised its angular aperture can be increased because the spherical aberration constant falls with the focal length. In magnetic lenses of present design, it results[18] that the power E delivered to the spot is proportional to r_2. As the camera length is also proportional to r_2, the intensity on the plate is proportional to r_2/r_2^2, or r_2^{-1}. In such conditions the exposure time would actually be reduced at higher resolution, but only so long as increasing the lens power reduced the focal length and hence r_2 also. With an object distance of 150 mm and with a minimum focal length of 1 mm, this limiting value of r_2 is 0.3μ. Any further reduction in spot size can be achieved only by reducing the size of the electron source or by increasing its distance from the lens, directly or by use of a condenser lens. The lens aperture must then be reduced, according to Eq. (2), to insure that spherical aberration gives a disk of confusion always less than the required resolution. Then $E \propto r_2^{8/3}$, and the reduction in exposure offered by reduced camera length is more than offset by the fall in incident energy. The intensity on the plate now depends on r^4 and the exposure time at optimum magnification is proportional to $1/r^4$. Improvement in resolution from 0.3μ to 0.1μ would require an increase in exposure of slightly more than four times, even with a reduction in camera length to one-third. Some compensatory increase in β may be obtained by raising the filament temperature, but only at the cost of shortening its life.

Two other points must be noted. First, the x-ray dosage of the object will not increase unduly as the resolution is raised; it varies to the same degree as does the exposure of the photographic emulsion.

Second, there is little to be gained by choosing a recording emulsion of greater sensitivity or higher resolution than average. Higher resolution can be obtained only by sacrificing sensitivity, so that although the emulsion may be used nearer the source, the exposure is not appreciably shortened. In any case, emulsions of very high resolution would have to be brought very close to the source of x-rays: if the grain size were 1μ, and the focal spot were 0.1μ, the emulsion would have to be no more than 1 mm from the spot for optimum use. We have preferred to use a lantern type of plate and a more convenient length of camera.

Stereography can be carried out very simply in this shadow projection method. All the specimen is in focus at once and with the same resolution, although different planes are magnified by different amounts. Three-dimensional imaging can thus be obtained by taking two photographs from slightly different viewpoints; as divergent illumination is employed, the specimen may either be tilted or translated between exposures. The latter method has been used, since it is simpler. The resulting stereographic pairs show directly the spatial configuration of a specimen in a way which is not possible in high power optical microscopy, owing to the limited depth of focus.

3. APPARATUS

The complete x-ray tube is shown in Figs. 3 and 4; it is mounted vertically to simplify the alignment of the parts. The electron gun emits a narrow beam, which is collimated by the condenser lens before reaching the objective, and this lens forms a focal spot on the target foil mounted on a washer of narrow bore to form a

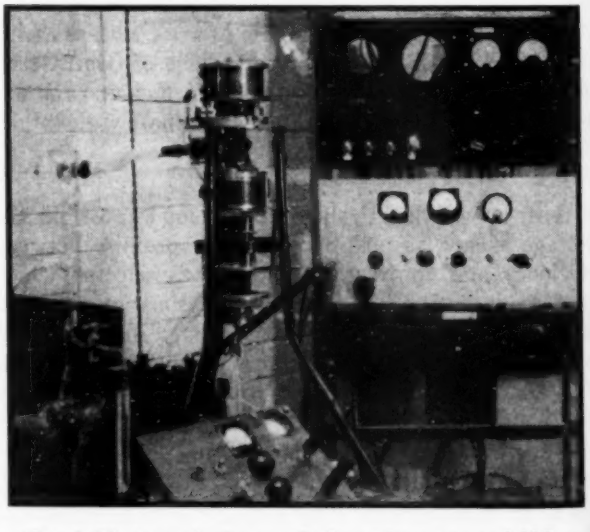


FIG. 4. Photograph of x-ray shadow microscope and 10-kv power supply.

Lenard window. The tube is mounted on a tripod to give mechanical stability, and is continuously evacuated by a mercury diffusion pump fitted with liquid air trap. The construction of the objective pole-pieces and the position of the specimen are shown in greater detail in Fig. 5.

Electron Gun

The electron beam is produced in a gun of triode type, as employed in electron microscopes. Wire of 50μ diameter is normally used for the V-shaped filament, but experiments have been made with 25μ wire. The centering of the filament tip in the shield aperture and its axial distance from it are critical if maximum brightness is to be obtained. The aperture in the shield is 1 mm and that in the anode 4 mm in diameter.

The anode voltage is generated electronically in a voltage doubler rectifier circuit operating at 100 kc. The output is variable between 3 and 13 kv, smoothed to 1 part in 2000. The total current drawn from the cathode is 60 μ-amp of which about 10 μ-amp is obtained in a spot of 1μ diameter at the target. The lenses are supplied with dc from a 120-v bank of batteries and the filament from 6-v accumulators in order to insure constancy.

Electron Lenses

The condenser lens has 3600 turns of copper wire and a bore of 19 mm. It is usually used to vary the effective distance of the electron source, without forming a real image of it before the objective lens. To minimize aberrations it is desirable to keep the latter at maximum strength and vary imaging conditions by means of the condenser. For attaining a still smaller focal spot pole-pieces could be fitted to the condenser, allowing a demagnified first image to be formed for subsequent further demagnification by the objective. An intermediate viewing screen can be inserted at S (Fig. 3) for observing the form of the beam entering the objective; it also assists preliminary rough alignment.

The objective lens first used was a modified version of an electron microscope objective. The pole-pieces had an internal bore of 6.3 mm and separation of 4.7 mm; at 1700 amp-turn excitation it gave a minimum focal length of 2.2 mm. Later, a projector lens of the original type EMB electron microscope was fitted with special pole-pieces, designed from the data of Liebmann and Grad,[13] having a bore of 3 mm and gap of 1.6 mm. It gives a minimum focal length slightly greater than 1 mm. With the condenser out of action the object distance is 380 mm, so that the maximum demagnification obtainable is 380 times. As the condenser is increased in power, the demagnification first decreases and is 150 when the electron beam is parallel to the axis between the lenses; later it increases again.

The arrangement of the target and specimen carrier in the objective is shown in Fig. 5. The electron beam is brought through the lens in a thin-walled vacuum liner. The target of tungsten foil is mounted on a disk having

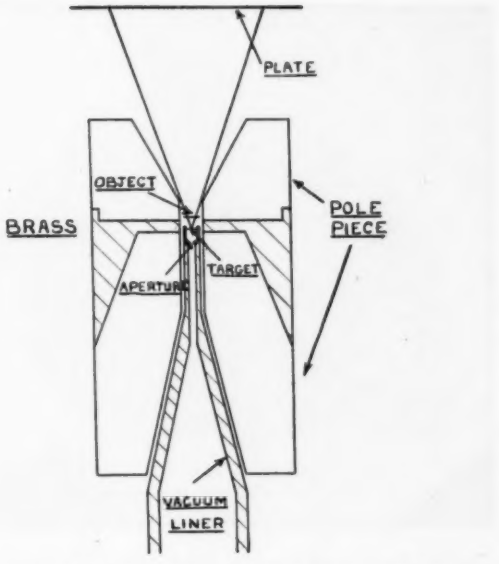


FIG. 5. Position of aperture, target, and object in objective lens pole pieces.

a central aperture 250μ in diameter, which is sealed to the end of the tube by Apiezon sealing compound. A smaller aperture within the tube limits the angular width of the beam reaching the target. The tube is mounted on a metal bellows which allows the target to be moved axially in the center of the lens until it is at the focal point. The specimen is mounted at the end of a short tube, which projects into the bore of the pole pieces and is fixed at its other end in a stage plate. Adjustments are provided for axial motion to bring the specimen close to the target, and for lateral motion to search the specimen. The shadowgraph is received on a photographic plate held in a light-tight tube (not shown) at a distance of a few centimeters above the target. To take stereographs, the specimen is translated slightly between successive exposures.

During preliminary alignment and focusing, the image of a fine grid is viewed on a fluorescent screen with a $\times 10$ eyepiece. At a voltage of 10 kv and with a spot size even less than 1μ, it is still possible to obtain enough x-ray intensity for visual focusing provided that the eye is first dark adapted. At lower voltages it becomes difficult to focus in this way; some preliminary tests have been made of the value of image intensifying tubes for overcoming this limitation.

The cone of rays received on the plate is defined by the tube carrying the specimen, which in turn is determined by the design of the pole pieces. Thus the field of view becomes progressively more restricted if the specimen is advanced towards the target, in order to obtain higher primary magnification as the spot size and hence the resolution is made finer. At the same time the camera length should be shortened, so that the resolved detail accords with the size of grain in the emulsion. The optimum exposure time thus remains approximately

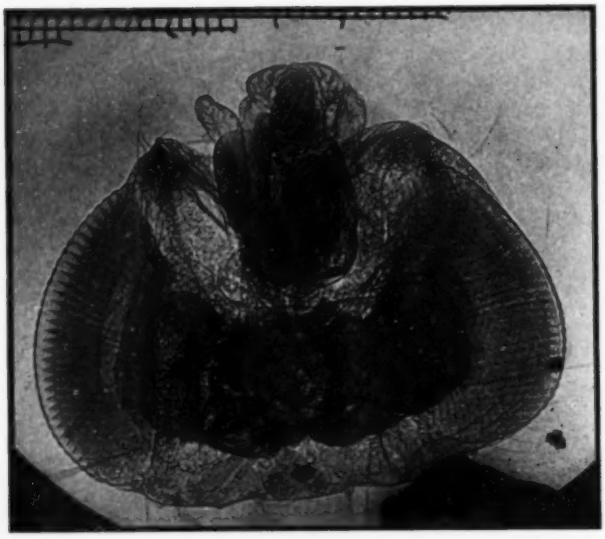


FIG. 6. Head of *Drosophila melanogaster*, freeze-dried. Original magnification ×30; reproduced at ×110. Repeat distance of reference grid = 17μ.

constant, and is of the order of 1 min. When a large field of view is desired, the specimen must be withdrawn from the target and the camera length correspondingly increased to maintain high magnification; the exposure time may then be of the order of 10 min.

FIG. 7. *Aphis fabae*, freeze-dried. Original magnification ×20; reproduced at ×100. Repeat distance of reference grid = 17μ.

4. RESULTS

The first step in taking x-ray shadow micrographs is to focus the electron beam by changing the objective lens current while viewing the image of a 1500 mesh/inch test grid on a fluorescent screen. The grid is then replaced by the specimen to be examined and the screen by a fine grain medium speed photographic plate, on which the exposure is made in the course of a few minutes. By taking a series of three or four exposures with varying values of the objective lens current the optimum condition for resolution can be found.

A few nonbiological specimens have been examined, but most of the results to date were obtained with insects of various species. Living specimens contain enough water to obscure the finer internal details, while dried specimens are liable to distortion both of individual features and of their mutual arrangement. Heavy metal injection with lead naphthenate gave some improvement in the cases of *Parthenothrips dracenae* and *Drosophila melanogaster* (the fruit fly).

The best results have been obtained by freeze drying *Drosophila melanogaster;* the specimens were frozen in liquid air and dried in vacuum at −30°C. All the water content was removed, and the internal detail then showed very clearly in the x-ray micrographs. Stereographic reproduction of freeze dried specimens has verified that very little distortion takes place in the process.

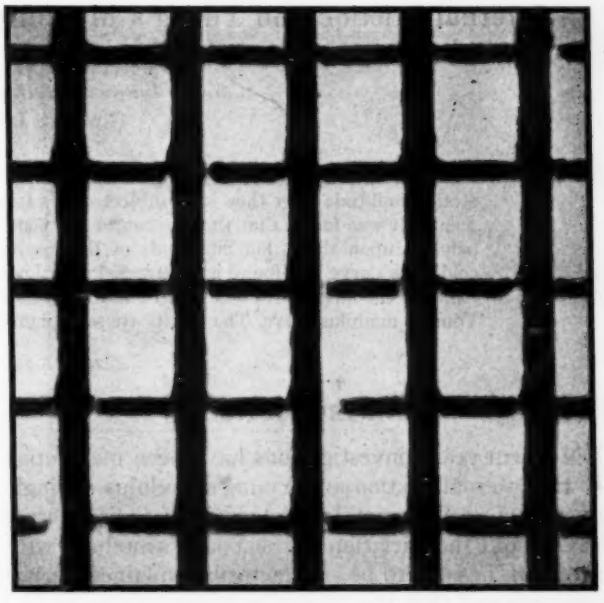
A micrograph of the isolated head of a fruit fly dried in this way is shown in Fig. 6, at a magnification of 110×. The dark area in the center of the head is the brain; the facets of the two compound eyes can be seen with the fibers running from each facet to the brain. A piece of 1500 mesh silver grid, with bars 3μ wide, is included in all micrographs for determining resolution and magnification. A freeze dried aphid (*Aphis fabae*) is shown in Fig. 7 at a lower magnification to include the whole specimen. Single cells can be seen in parts of the abdomen.

As a direct comparison of the new technique with optical microscopy, an x-ray micrograph of the 1500-mesh grid is reproduced in Fig. 8 at 930×, alongside the best obtainable photomicrograph. The latter appears much sharper, owing to diffraction effects at the edges. These are absent in the x-ray case, and in addition the grid bars are thinned towards the edge sufficiently for the x-rays to penetrate them; as a result, the grid shows poorer contrast as well as being less sharp. However, careful examination of the detail visible in the imperfections of the grid shows that the resolution in the shadowgraph is little worse than that in the optical picture. A resolution of about 0.5μ can be claimed on the basis of this evidence, and a comparable figure is indicated by the fine detail shown in the head of the fruit fly (Fig. 6).

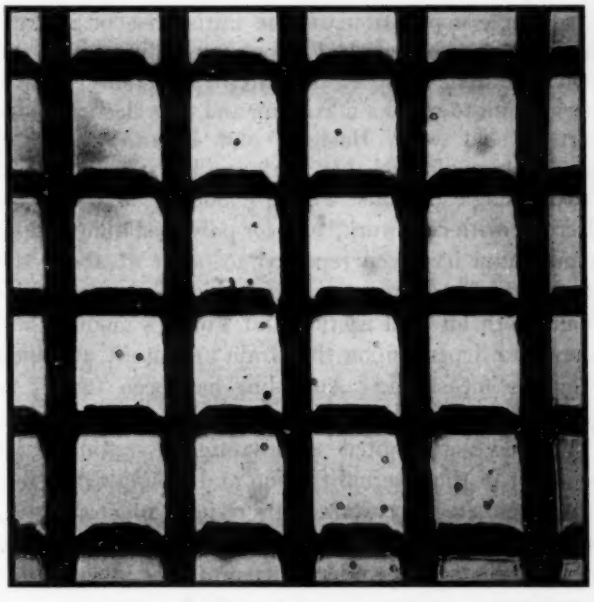
5. CONCLUSION

The experiments so far made support the theoretical indications that x-ray shadow micrography is a practicable technique at least up to the resolution of the optical microscope. The chief limitations on the performance of the present apparatus are probably the astigmatism of the objective lens and the blurring of the focal spot in the target foil. There should be no great difficulty in improving it in these respects so as to obtain a resolution of 0.1μ. By shortening the camera and possibly overrunning the filament still more, microradiography at the resolution of the ultraviolet microscope should be possible with an exposure time of the order of 10 min.

Greater difficulties stand in the way of achieving a resolution of 100A, which can be regarded as the practical limit of fineness of an electron beam. The x-ray output from such a small spot, using electron beam intensities of the order at present known, would be so weak that exposure times would rise to hours. If spherical aberration in electron lenses could be corrected, it would be possible to use much larger apertures and thus to increase greatly the input of electron energy to the target; the rate of dissipation from such minute spots is rapid enough to permit such an increased input. In the absence of lens correction the only available means of reducing the exposure time would be to use electron sources of much higher specific emission. Marton[21] has proposed the use of pulsed cold emission from a tungsten point, and this method deserves experimental exploration in the circumstances of x-ray microscopy. As an alternative, the development of image intensifiers for x-rays[22] may ultimately offer a means of circumventing the limitation set by initial intensity, but it is doubtful if the necessary increase in brightness can be obtained without loss of resolution. It may prove that the practical limit of the shadow projection technique is little below 1000A, but this would appear to be considerably better than that obtainable by the contact or mirror methods of magnifying objects with x-rays.



(a)



(b)

FIG. 8. (a) X-ray micrograph of 1500 mesh per inch silver grid with 3μ bars. Original magnification ×100; reproduced at ×930. (b) Optical micrograph of similar grid, taken with oil immersion objective at $N.A.=1.3$; reproduced at ×930.

ACKNOWLEDGMENTS

We wish to express our thanks for assistance in the development of this work to the Universities of London and Oxford (V.E.C., 1939–41), to the University of Cambridge (V.E.C., since 1946), to the Research Council of Ontario and the Department of Veterans Affairs of Canada for grants (W.C.N.). For the gift of materials and apparatus we are indebted to Electrical and Musical Industries, Ltd., Metropolitan-Vickers Ltd., and the Philips Company, Eindhoven, Holland. For provision of specimens and discussion of the biological problems, we are particularly indebted to Professor V. B. Wigglesworth and Dr. J. R. G. Bradfield.

[21] L. Marton, Symposium on Electron Physics, Washington (1951).

[22] W. S. Lusby, Elec. Eng. **70**, 292 (1951).

Internal Friction and Young's Modulus of Cold-Worked Copper Single Crystals*

J. WEERTMAN† AND J. S. KOEHLER‡
Carnegie Institute of Technology, Pittsburgh, Pennsylvania
(Received December 31, 1952)

The internal friction and Young's modulus of single crystals of copper were measured as a function of strain amplitude after they were subjected to various amounts of cold work and different annealing treatments. It was found that the position of the peak in the internal friction *versus* cold-work curve was dependent upon the strain amplitude of the measurement. No minimum in the Young's modulus *versus* cold-work curve was found as contrasted with Lawson's result for polycrystalline copper. A minimum was found in the internal friction *versus* annealing time curve, and a maximum was found in the corresponding Young's modulus curve. The results are semiquantitatively explained using a dislocation model.

I. INTRODUCTION

IN recent years, investigations have been made upon the internal friction and Young's modulus of single crystals of various metals. The origin of their internal friction and the variation in the Young's modulus with cold work has yet to be satisfactorily explained. It has been assumed that the origin lies in the motion of dislocations. Since cold work is supposed to increase the density of dislocations, this assumption would give a qualitative explanation of the initial rise of internal friction with cold work. For polycrystalline brass[1] and polycrystalline copper[2] the internal friction has also been found to reach a maximum and then decrease with further cold work. Hasiguti and Hirai[3] have independently confirmed that the same behavior occurs in copper single crystals. Young's modulus is known to decrease with cold work,[4] and for polycrystalline copper a minimum has been reported to occur at about the strain which produces a maximum in the internal friction.[2] Both internal friction and Young's modulus are known to depend upon the strain amplitude at which they are measured.[4,5] Annealing has been found, in general, to have the opposite effect of cold work.[5]

We have attempted to investigate carefully the behavior of the internal friction and Young's modulus of copper single crystals with respect to cold work, annealing treatment, and strain amplitude of the measurements, and we have attempted to give a semiquantitative explanation of our results in terms of a dislocation model.

* This work was supported by the U. S. Office of Naval Research.
† Abstract of thesis submitted June, 1951 to Carnegie Institute of Technology in partial fulfillment for the requirements of the degree of Doctor of Science. Present address, Naval Research Laboratory, Washington, D. C.
‡ Present address, Department of Physics, University of Illinois, Urbana, Illinois.
[1] Zener, Clarke, and Smith, Trans. Am. Inst. Mining Met. Engrs. 147, 90 (1942).
[2] A. W. Lawson, Phys. Rev. 60, 330 (1941).
[3] R. R. Hasiguti and T. Hirai, J. Appl. Phys. 22 1084 (1951).
[4] T. A. Read, Phys. Rev. 58, 371 (1940). T. A. Read, Trans. Am. Inst. Mining Met. Engrs. 143, 30 (1941).
[5] J. Marx and J. S. Koehler, Symposium on Plastic Deformation of Crystalline Solids, U. S. Office of Naval Research, Pittsburgh (1950), p. 171.

II. EXPERIMENTAL

The internal friction and Young's modulus were measured, using the modification of the piezoelectric method introduced by Marx.[6] The experimental arrangement is shown in Fig. 1. Specimens were glued with beeswax to the bottom of the driver crystal, which was an 18.5°, X-cut quartz crystal silvered on two opposite faces. On the top of the driver crystal was glued another quartz crystal, the gauge crystal. The fundamental frequencies of free longitudinal vibration of the driver and the gauge crystal differed by less than 0.1 percent. The fundamental frequency of the specimens differed by less than 5 percent from those of the driver. All measurements were made near the frequency 38.5 kc. Measurements were made by slowly changing the frequency of a constant applied voltage to the driver crystal until the voltage across the gauge crystal reached a maximum. Marx[6] has shown that the decrement§ of the composite oscillator is equal to a constant times the ratio of the driver voltage divided by the maximum gauge voltage. The maximum strain amplitude in the specimen is proportional to the gauge voltage. Both constants can be evaluated by measuring the internal friction by the ac bridge method.[4] The decrement of the specimen alone can be calculated from that of the composite oscillator by the equation[4]

$$\Delta_c m_c = \Delta_s m_s + \Delta_g m_g + \Delta_d m_d, \tag{1}$$

where Δ_i and m_i stand for the decrement and mass of the composite oscillator, specimen, driver, and gauge crystal.

Young's modulus of the specimen was calculated from the equation $E=\rho(2lf_s)^2$, where ρ is the density, l is the length of the specimen, f_s is the fundamental resonant frequency of free longitudinal vibration, and E stands for Young's modulus. The value of f_s was calculated from the resonant frequency of the composite oscillator by the equation

$$f_c m_c = f_s m_s + f_d m_d + f_g m_g, \tag{2}$$

[6] J. Marx, Rev. Sci. Instr. 22, 503 (1951).
§ The decrement is defined to be equal to the energy lost per cycle in an oscillating system divided by twice the stored energy in the system. It is equal to the logarithmic decrement if it is small and independent of strain amplitude.

where f_i stands for the resonant frequency of the composite oscillator, driver, etc.

Specimens

Foot long crystals of 0.38-in. square cross section were grown in vacuum by the Bridgman method from 99.999 percent pure copper supplied by the American Smelting and Refining Company. After being oriented by means of x-rays, the specimens were embedded in micarta tubes with plaster of Paris. Specimens of the proper length were then cut from the tubes containing crystals and were ground down to such a length that they would have a resonant vibrational frequency near

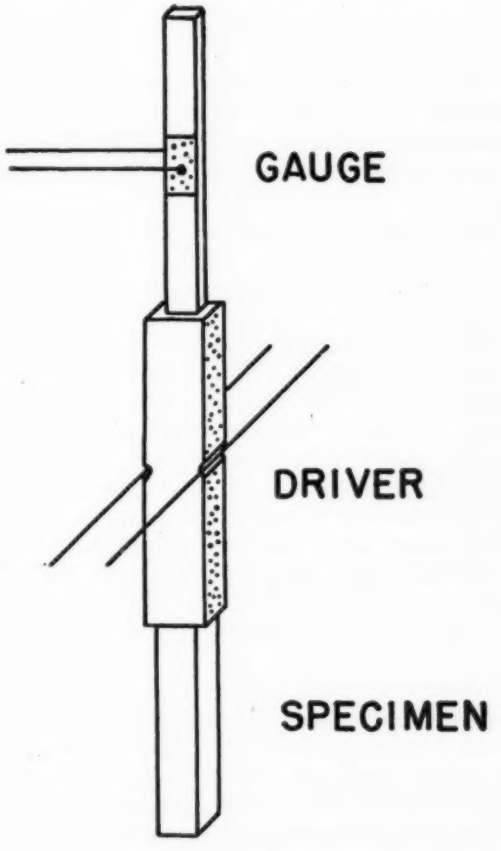


FIG. 1.

38 556 cps. Their approximate length was 1.5 in. After the worst of the cold-worked surfaces were etched away, the specimens were annealed in vacuum at around 950°C for about nine hours. Several specimens were annealed in dry hydrogen. After annealing, the specimens were cold worked by subjecting them to a compression perpendicular to their long axis. The compression was applied using a double piston arrangement which is shown in Fig. 2. After a load was applied, the internal friction and the modulus were measured. Then the specimen was replaced in the compression device and a larger load applied. The load was left on for a period of three minutes.

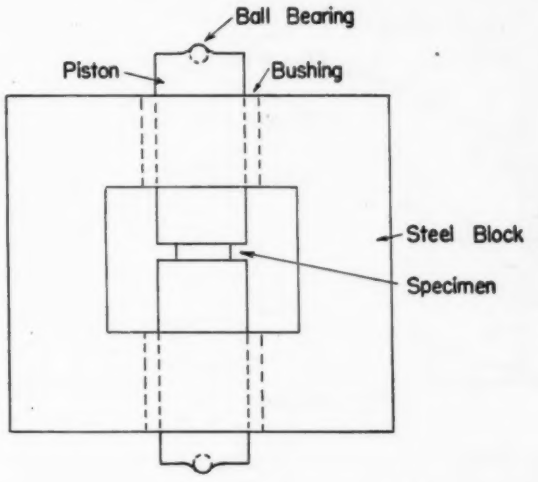


FIG. 2.

III. RESULTS

In Figs. 3 and 4 typical curves of the internal friction *versus* strain amplitude are shown after various amounts of cold work. Figures 5 and 6 show the strain amplitude dependence of Young's modulus after different amounts of cold work. In Figs. 7 and 8 typical cases of the decrement and Young's modulus at a constant strain amplitude as a function of cold work are given. Since deforming the specimen changes its orientation relative to the long axis of the quartz driver, the measured modulus should be corrected to give the modulus of the original orientation. In Fig. 7 the smooth curve shows the measured modulus. The dotted curve shows the maximum possible correction that could be made to the measured modulus. The actual modulus would lie between these two curves.

By plotting a series of curves, each for a different constant strain amplitude, of internal friction *versus* amount of cold work, such as is shown in Fig. 7, one can obtain the dependence of the amount of cold work at maximum decrement on the strain amplitude. This dependence is shown in Fig. 9, where the strain amplitude is plotted against the resolved shear stress required

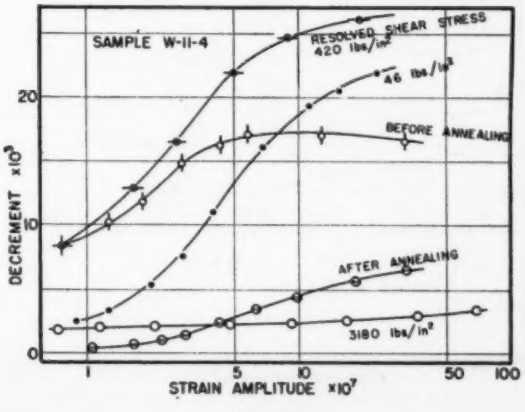


FIG. 3.

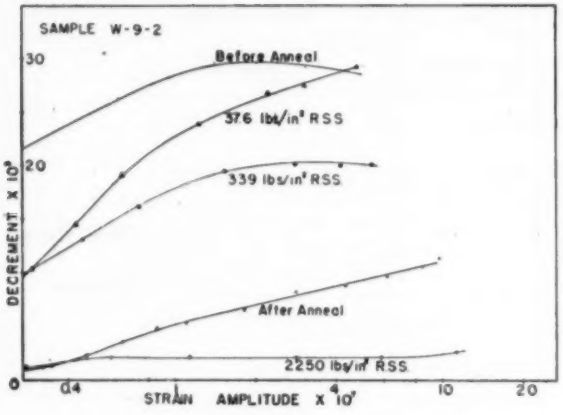


FIG. 4.

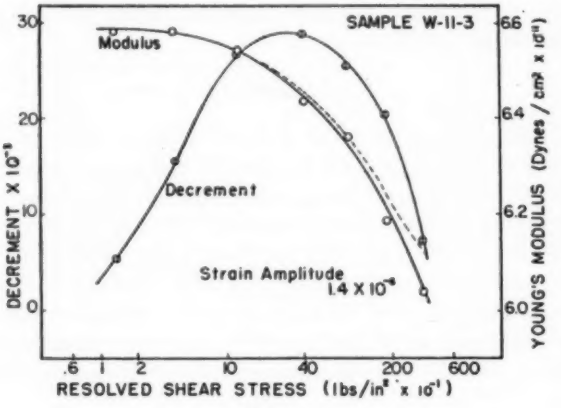


FIG. 7.

to obtain the maximum decrement. Also in Fig. 10 is plotted the maximum obtainable internal friction at different strain amplitudes. In Fig. 11 is plotted the change of internal friction at a constant strain amplitude as a function of annealing time. The temperatures refer to the temperature of the vacuum anneal preceding the measurement. The initial value of the decrement in Fig. 11 is that obtained after the sample had been extensively cold worked. Young's modulus changes upon annealing as shown in Fig. 12 where E_i represents the value of the modulus after a complete anneal at 1000°C.

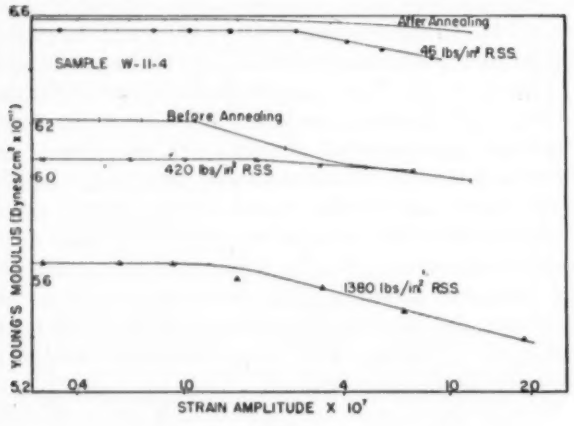


FIG. 5.

When a specimen is cold worked again after it has been extensively cold worked and then annealed at high temperatures, one can again obtain a maximum in the cold work *versus* decrement curve. However, for a given strain amplitude, the maximum obtainable decrement is smaller, and it takes a larger amount of cold work to reach the maximum decrement.

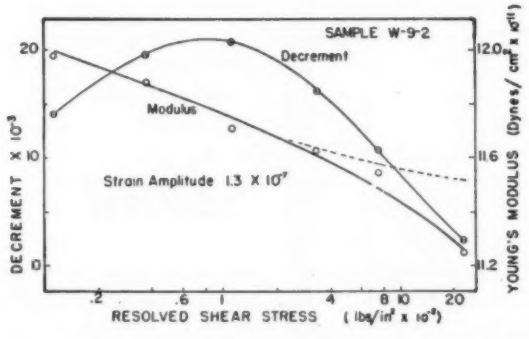


FIG. 8.

In Table I is given a summary of the results on the

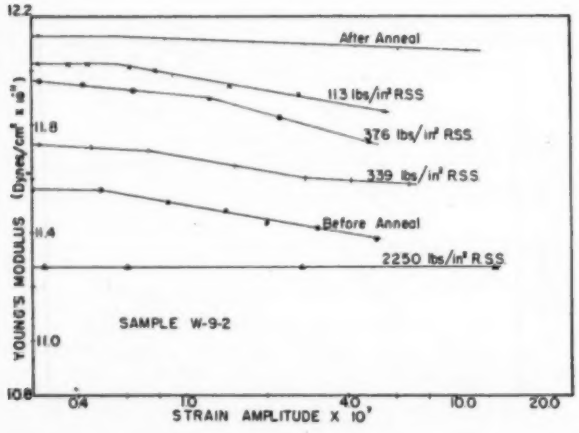


FIG. 6.

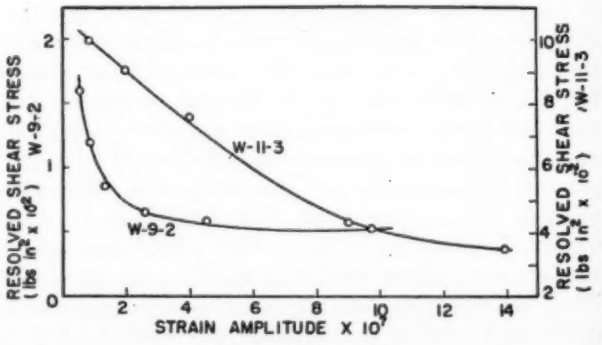


FIG. 9.

TABLE I.

Specimen	Crystal	r.s.s. at Δ_{max}[a] ($\epsilon_0 = 8\times10^{-5}$)	r.s.s. at Δ_{max} ($\epsilon_0 = 2\times10^{-5}$)	Remarks
W-3	W-3	less than 365[c]	less than 365	vacuum annealed[b]
W-8-2	W-8	less than 150	less than 150	vacuum annealed
W-8-3	W-8	less than 100	less than 100	vacuum annealed
W-9-1	W-9	100	75	vacuum annealed
W-9-2	W-9	130	50	vacuum annealed
W-11-2	W-11	285	285	annealed in bad vacuum
W-11-3	W-11	1100	300	hydrogen anneal
W-11-4	W-11	520	190	vacuum annealed
W-11-5	W-11	1500	900	hydrogen anneal

Specimen	Δ_{max} ($\epsilon_0 = 8\times10^{-5}$)	Δ_{max} ($\epsilon_0 = 2\times10^{-5}$)
W-8-2	0.020	0.029
W-8-3	0.021	0.030
W-9-1	0.013	0.034
W-9-2	0.013	0.034
W-11-2	0.0075	0.012
W-11-3	0.015	0.033
W-11-4	0.010	0.029
W-11-5	0.013	0.028

[a] Resolved shear stress (in lb/in.^{2}) at maximum decrement. ϵ_0 is the strain amplitude.
[b] The vacuum was better than 10^{-4} mm of mercury.
[c] When it is stated that the resolved shear stress for maximum decrement is *less than* some number, it means that the maximum in the decrement had been passed at the first load used.

compression specimens. Table II gives crystal orientation data.

Measurements were also made upon samples cold worked in tension. A pair of tensile specimens similar to those of Neurath and Koehler[7] was grown as one single crystal. One of the pair would be pulled with a load of about 3000 lb/in.^{2} r.s.s. The other was either not pulled or pulled with a load of about 103 lb/in.^{2} r.s.s. This procedure was carried out with seven pairs of specimens, three pairs of which had been given an anneal before the pulling. After the specimens were pulled, the grips were sawed off, and the specimens were ground and etched to their proper length.

TABLE II.

The orientations of the specimens before loading can be given as follows:

X = angle between direction of compression and the active glide plane.
λ = angle between direction of compression and the active slip direction.
θ = angle between long axis of specimen and (100) axis.
ϕ = angle between long axis of specimen and (010) axis.
ψ = angle between long axis of specimen and (001) axis.

For crystals 9 and 11 these angles were:

Crystal 9, i.e., W-9-1 W-9-2	$X = 46°$	
	$\lambda = 45°$	
	$\theta = 68°$	
	$\phi = 32°$	
	$\psi = 70°$	
Crystal 11, i.e., W-11-2 W-11-3 W-11-4 W-11-5	$X = 35°$	
	$\lambda = 55°$	
	$\theta = 0°$	
	$\phi = 90°$	
	$\psi = 90°$	

[7] P. Neurath and J. S. Koehler, J. Appl. Phys. **22**, 621 (1951).

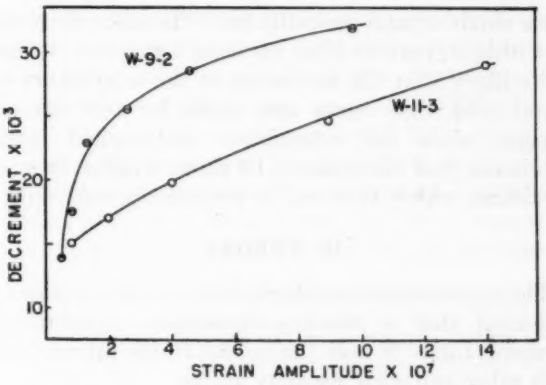


FIG. 10.

In all seven cases the extensively cold-worked specimens had much smaller decrements than their mates. In addition, the decrements of the highly cold-worked samples were practically independent of strain amplitude. The decrements of these samples correspond to that of the 3180 lb/in.^{2} curve of Fig. 3. The specimens

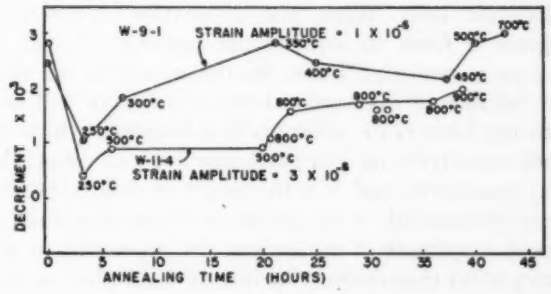


FIG. 11.

which were not pulled, or only pulled by a small load, always had a very high decrement corresponding to the 420 lb/in.^{2} curve of Fig. 3. Their decrement was always very dependent on the strain amplitude.

Since compression or tension specimens cut from the

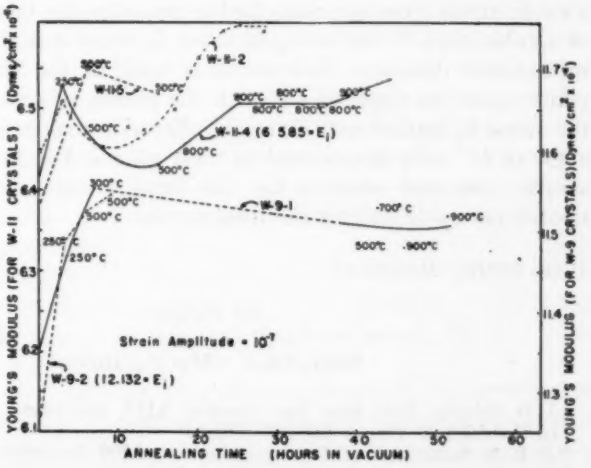


FIG. 12.

same single crystal generally have the same decrement to within 50 percent after identical treatment, it seems quite likely that the maximum in the internal friction *versus* cold work curve also exists for cold work by tension, since the extensively cold-worked tension specimens had decrements 10 times smaller than the specimens which were not so extensively cold worked.

IV. THEORY

The experiments described above can be explained by assuming that a moving dislocation experiences a damping force[8–10] and that dislocations interact with each other and with impurity atoms.

A complete theory would treat the small oscillations of impurity-anchored dislocations[11] which are present in such concentrations that they interact with one another. However, the complete calculation does not as yet exist. We have made calculations using a simplified model in which the displacement of a dislocation element is independent of its position along the length of the dislocation, i.e., we deal essentially with the average displacement of the various elements along a dislocation loop. When the dislocation is moved a distance x from its equilibrium position, it will experience a restoring force. In the appendix we show that because of dislocation interaction there will be a restoring force GNx, where G is a constant which depends sensitively on the particular type of dislocation array considered and N is the length of dislocation line per cc of material. If we assume for simplicity that for a given amplitude of oscillation, the impurities or any agency other than dislocation interactions cause another restoring force of G_0x, where G_0 is a constant, then the equation of motion for a unit length of dislocation can be written as

$$M(d^2x/dt^2)+B(dx/dt)+(GN+G_0)x=F\cos(wt), \quad (3)$$

where M is the effective mass per unit length of dislocation, $B(dx/dt)$ is the dissipative force per unit length, and $F\cos(wt)$ is the applied force on the dislocation. If a cyclic stress $\sigma\cos(wt)$ is applied to two opposite faces of a cube, then F will be equal to $\sigma a/2$, where a is the interatomic distance. This result is valid if the slip plane makes an angle of 45° with the planes to which the stress is applied and if the slip direction makes an angle of 45° with the normal to these planes. For this simple case one obtains for the total energy dissipated per cycle and for the total strain:

$$\text{Total energy dissipated} = \frac{\pi N\sigma^2a^2Bw}{16(G_0+GN-Mw^2)^2+16B^2w^2}, \quad (4)$$

$$\text{Total strain} = \frac{\sigma\cos(wt)}{E} + \frac{N\sigma a^2[(G_0+GN-Mw^2)\cos(wt)+Bw\sin(wt)]}{4\sqrt{2}[(G_0+GN-Mw^2)^2+B^2w^2]}, \quad (5)$$

where E is the true Young's modulus for the material if there were no dislocations. The decrement will be given by the equation

$$\Delta=\frac{(\pi/16)Ea^2NBw}{(G_0+GN-Mw^2)^2+B^2w^2}. \quad (6)$$

The effective modulus will be given by the ratio of the value of the stress when the strain is at its maximum divided by the maximum strain. If $\bar{E}$ is the effective modulus and E the modulus if no dislocations were present, then we would get the equation

$$\frac{E-\bar{E}}{E}=\frac{\sqrt{2}(ENa^2)(G_0+GN-Mw^2)}{8[(G_0+GN-Mw^2)^2+B^2w^2]}+\frac{2E^2N^2a^4B^2w^2}{64[(G_0+GN-Mw^2)^2+B^2w^2]^2}. \quad (7)$$

Frank[12] has shown that the effective mass per unit length is of the order of the density of the material times the square of the interatomic distance. For the frequency we used, Mw^2 is much smaller than the term GN, if we use the calculated value of G and assume that N must be larger than 10^5. We will therefore neglect the Mw^2 term in the following discussion. This essentially says that the inertial force on the dislocation is unimportant.

Let us first consider the behavior of the decrement. From Eq. (6) we see that the decrement will increase as N increases, then reach a maximum, and finally decrease. Since cold work is supposed to increase the value of N, this equation gives a qualitative explanation of the observed maximum in the internal friction *versus* cold work curve.

The dislocation density N_m at maximum decrement obtained from Eq. (6) is

$$G^2N_m^2=G_0^2+B^2w^2. \quad (8)$$

The value of the maximum decrement obtained using condition (8) is

$$\Delta_{max}=\frac{\pi Ea^2N_mBw}{32(G_0GN_m+G^2N_m^2)}. \quad (9)$$

Since at the higher values of strain amplitude the dislocations are more likely to break away from impurities, the effective value of G_0 would be smaller at the higher strain amplitudes. From Eq. (6) this would

[8] J. D. Eshelby, Proc. Roy. Soc. (London) **A197**, 396 (1949).
[9] G. Leibfried, Z. Physik **127**, 344 (1950).
[10] F. R. N. Nabarro, Proc. Roy. Soc. (London) **A209**, 278 (1951).
[11] J. S. Koehler, *Imperfections in Almost Perfect Crystals* (John Wiley and Sons, Inc., New York, 1952), p. 197.
[12] F. C. Frank (Physical Society, London 1948), p. 46.

mean qualitatively that the decrement would increase with increasing strain amplitude unless the specimen was cold worked so much that GN was much larger than G_0. For this latter case the decrement would be independent of strain amplitude. Also, from Eq. (8) and from our information about G_0, we find that the maximum in the internal friction would occur for a smaller amount of cold work at the higher strain amplitudes since a smaller value of N_m is required. From Eq. (9) we find that the value of the decrement at its maximum is smaller at the lower strain amplitudes since the value of G_0 and hence N_m is larger than at the higher amplitudes. Thus Eqs. (6), (8), and (9) can be used to give a qualitative explanation of the behavior of the decrement.

Let us next consider the behavior of Young's modulus by using Eq. (7). On varying N, the first term on the right does not contain a maximum, but the second term does. This maximum occurs when N equals N_m, i.e., the value which makes the decrement a maximum. If we substitute Eq. (8) into Eq. (7), we obtain

$$\left(\frac{E-\bar{E}}{E}\right)_{\Delta=\max}=\frac{\sqrt{2}Ea^2}{16G}+\left(\frac{\sqrt{2}Ea^2}{16G}\right)^2\left(\frac{GN_m-G_0}{GN_m+G_0}\right)^2. \quad (10)$$

It can be seen that even when the second term on the right of Eq. (7) takes its maximum value, it can be neglected in comparison with the first term. Note that this conclusion does not depend on the numerical values of the quantities involved. Equation (7) thus predicts that the effective modulus will continue to decrease as N is increased, any increase resulting from the second term being insignificant.

Now let us consider how well Eqs. (9) and (10) agree quantitatively with experimental values. Fortunately Eq. (10) is independent of both N and B, two highly uncertain quantities, and only depends on G, which should be correct at least as to an order of magnitude. Using the calculated value of G for the case in which local regions have equal numbers of positive and negative dislocations, the value of $(E-\bar{E})/E$ is 0.045 as compared to an experimental value of 0.035 when the decrement is at its maximum.

If Bw is much larger than G_0, then the expression for the maximum decrement becomes

$$\Delta_{\max}=\frac{\pi Ea^2}{32G}. \quad (11)$$

Using the calculated value of G, this expression is equal to about 0.05. The largest values of the decrement measured experimentally were of the order of 0.03. Although the numerical values agree reasonably well, one must still conclude that G_0 plays some role in determining the decrement, since Fig. 10 shows that Δ_m depends on strain amplitude, even when it takes these large values, and we have already seen that the amplitude dependence of G_0 determines that of Δ_m. However, if we consider that the value of N is 10^8 at the maximum decrement, then B is equal to 0.3 dyne-sec/cm^2, which is a value 10^4 larger than the theories of Eshelby or Leibfried predict. It seems more likely then that G_0 is much larger than Bw. For this case, Eq. (9) becomes

$$\Delta_{\max}=\frac{\pi Ea^2Bw}{64G^2N_m}=\frac{\pi Ea^2Bw}{64G_0G}. \quad (12)$$

Since B and N are rather uncertain, substituting values into Eq. (12) is not too meaningful. However, for a value of N equal to 10^7 and B equal to 5×10^{-3} dyne-sec/cm^2, we get the right order of magnitude for the maximum decrement. This value of B is also required to explain the decrement data of annealed copper crystals.[11]

The annealing behavior can now be examined. Since $G_0 \gg Bw$ the Eqs. (6) and (7) can be simplified to

$$\Delta=\frac{(\pi/16)Ea^2NBw}{(G_0+GN)^2}, \quad (13)$$

and

$$\left(\frac{E-\bar{E}}{E}\right)=\frac{\sqrt{2}Ea^2N}{8(G_0+GN)}, \quad (14)$$

where the second term in Eq. (7) has been neglected. The low temperature vacuum anneals produce a minimum in both Δ and dE/E. Since this phenomenon occurs in both quantities, it appears reasonable to ascribe it to some process or processes which at first increase G_0 or G or both and then decrease them. For example, it could be that the vacancies and interstitial atoms produced during deformation diffuse to the dislocations anchoring them, thus producing the initial drop by increasing G_0. The subsequent rise could occur as the result of the migration along the dislocations, in which case by recombination or by agglomeration the number of pinning points, and hence, G_0 would be reduced. The subsequent rise in Δ and drop in dE/E are probably produced by a drop in the number N of dislocations. The signs of the slopes predicted on this basis by Eqs. (13) and (14) are in agreement with experiment providing $GN > G_0$, i.e., provided one does not succeed in teaching a value of N which corresponds to a maximum in Δ. Since a repetition of the entire experiment after a "complete" anneal produces a $\Delta_{\max}$ which is smaller than that obtained on the initial cold working, we must conclude that G_0G is increased during a cold-work, annealing cycle.

It should be admitted that the explanations proposed for the annealing behavior are tentative and that additional research on the effects produced by annealing would be welcome.

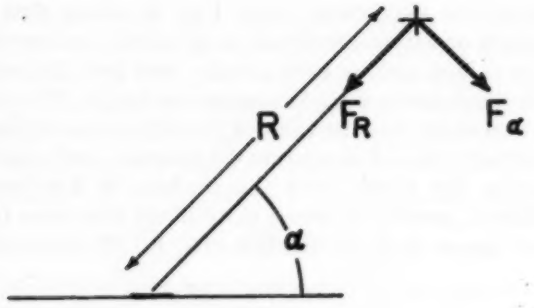


FIG. 13.

We would like to acknowledge the aid and guidance of Dr. John Marx, who assisted materially in the planning of the experiment. The first author would also like to thank Dr. J. E. Goldman and Dr. R. Smoluchowski for valuable discussions regarding the theory.

APPENDIX

Consider the Taylor dislocations given in Fig. 13. The negative dislocation exerts a force[13] on the positive one whose components are

$$F_R=C/R,$$

$$F_\alpha=\frac{C\sin 2\alpha}{R}, \tag{15}$$

where C is equal to 6×10^{-5} dyne for copper. Consider the dislocation array in Fig. 14. On application of a shearing stress, the odd rows move a distance x to the left and the even rows the same distance to the right. To the first approximation the restoring force exerted on a positive dislocation by four negative dislocations, as is shown in Fig. 15, will be

$$\text{force}=\frac{8C}{(n^2+m^2)R^2}\left[x-\frac{8n^2m^2x}{n^2+m^2}\right]. \tag{16}$$

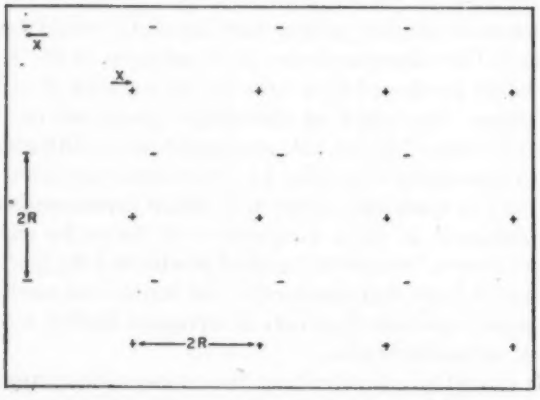


FIG. 14.

[13] A. Cottrell, *Progress in Metal Physics* (Interscience Publishers, Inc., New York, 1949), Vol. 1, p. 86.

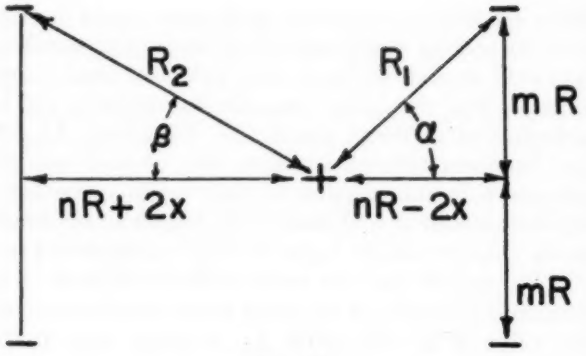


FIG. 15.

The total restoring force will be given by

$$\sum_{n,m}\frac{8Cx}{R^2}\left[\frac{1}{n^2+m^2}-\frac{8n^2m^2x}{n^2+m^2}\right], \tag{17}$$

where $n=1,3,5\cdots(N/2)^{\frac{1}{2}}$; $m=1,3,5\cdots(N/2)^{\frac{1}{2}}$. For large values of n and m the above sum can be replaced by an integral of the form

$$\int_0^{\pi/2}\int_{R_1}^{R_2}\frac{(1-8\cos^2\theta\sin^2\theta)}{R^2}RdRd\theta, \tag{18}$$

which is equal to zero. Therefore, it is necessary to use only the first few terms of Eq. (18) for an evaluation of the force. Using values of n and m equal to 1, 3, and 5, we get that the total force is equal to $-(0.76)8Cx/R^2$. If N is equal to the number of dislocations crossing one sq cm of material, then R^2 is equal to $1/2N$, and the total force is equal to $-GNx$, where G is equal to $12.2C=7.3\times10^{-4}$ dyne in the case of copper.

For the case in which the dislocations are arranged as shown in Fig. 16, one finds by exactly similar methods that

$$G=\frac{2bB\left\{a+\dfrac{a'-a}{2j-1}\right\}}{\left[(\tfrac{2}{3}j-1)(j-1)+\dfrac{B}{B'}(2j-1)\right]},$$

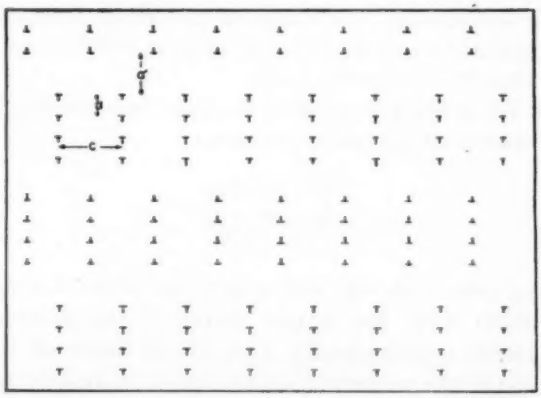


FIG. 16.

where

$$N=\frac{1}{c\left\{a+\dfrac{a'-a}{2j-1}\right\}},$$

$$B=\frac{\pi\mu b^2}{c^2(1-\nu)\left\{\cosh\left(\dfrac{2\pi a}{c}\right)-1\right\}}\times\left[\frac{\dfrac{2\pi a}{c}\sinh\left(\dfrac{2\pi a}{c}\right)}{\left\{\cosh\left(\dfrac{2\pi a}{c}\right)-1\right\}}-1\right].$$

In the above expressions, μ is the modulus of rigidity, b is the smallest interatomic distance, and ν is Poisson's ratio. To obtain B', one uses the expression for B, but a is replaced by a'. This second arrangement may find application in dealing with the dislocations produced during slip. In aluminum Brown[14] and Heidenreich and Shockley[15] have found a surface lamella structure which may indicate that the resulting internal dislocation pattern is of the type used in Fig. 13. Such would probably be the case if each group of slip lamellas, i.e., each slip zone, resulted from the action of a single Frank-Read[16] dislocation source. In any event, it is clear from the expressions for G in the two cases treated that the restoring force produced by dislocation interaction is strongly dependent on the type of dislocation array considered. This is fortunate, since it means that decrement measurements can give information on the particular kinds of dislocation patterns found in crystals.

[14] A. F. Brown, J. Inst. Metals **80**, 1331 (1951).
[15] R. D. Heidenreich and W. Shockley, *Report of a Conference on the Strength of Solids* (Physical Society of London, 1948), p. 57.
[16] F. C. Frank and W. T. Read, Phys. Rev. **79**, 722 (1950).

JOURNAL OF APPLIED PHYSICS VOLUME 24, NUMBER 5 MAY, 1953

Double Activators in Oxide Cathode Base Alloys*

A. Eisenstein, H. John, and J. H. Affleck
Department of Physics, University of Missouri, Columbia, Missouri
(Received December 31, 1952)

The presence of tungsten as a major impurity in the base metal of an oxide-coated cathode leads to the formation of the interface Ba_3WO_6. If silicon is also present, but as a minor impurity, this element gradually replaces tungsten in the interface to form Ba_2SiO_4. Similar reactions between titanium and aluminum with other known interface compounds were examined.

INTRODUCTION

FOR many years oxide-coated cathodes have been prepared on a nickel base metal containing small amounts of intentionally added impurities. These impurities were probably first added to produce a base alloy having good mechanical properties, but it was soon found that these impurities also affect the electrical characteristics of the cathode.[1] When reducing impurities such as Si, Ti, Al, or W are present, solid state chemical reactions occur in the cathode forming the interface compounds Ba_2SiO_4, Ba_2TiO_4, $BaAl_2O_4$, and Ba_3WO_6, respectively.[2–4] Other impurities are known to react with the oxide coating, but the identification of the interface is not certain. The presence of an interface compound indicates that a chemical reaction has taken place and all of these in which the metallic impurity combines with BaO release free barium. This is believed to aid the actviation of the cathode; thus the reaction may be considered as a beneficial process. However, some interface compounds, e.g. Ba_2SiO_4 and $BaAl_2O_4$, have a high specific electrical resistivity[2] and under some conditions of cathode operation this resistivity leads to deleterious effects.[5–7]

When a base alloy of 4.7 percent tungsten-nickel† is used, cathodes of superior thermionic activity are produced[8,9] which have long life. These cathodes do not exhibit a measurable interface resistivity[10] when life tested under conditions which give rise to an objectionable resistivity in silicon-nickel base cathodes. Following the completion of a 6000-hour life test of tungsten-nickel base cathodes, x-ray diffraction studies of the

* Supported in part by the U. S. Office of Naval Research.
[1] E. M. Wise, Proc. Inst. Radio Engrs. **25**, 714 (1937).
[2] H. P. Rooksby, Nature **159**, 609 (1947).
[3] D. A. Wright, Proc. Phys. Soc. (London) **62**, 188 (1949).
[4] E. G. Steward and H. P. Rooksby, Acta Cryst. **4**, 503 (1951).
[5] A. Eisenstein, J. Appl. Phys. **20**, 776 (1949).
[6] A. Eisenstein, J. Appl. Phys. **22**, 138 (1951).
[7] W. W. Lindemann and H. J. Hannam, Bull. Am. Phys. Soc. **27**, 5 (1952).
† Obtained from E. M. Wise, International Nickel Company.
[8] H. A. Pohler, Proc. Inst. Radio Engrs. **40**, 190 (1952).
[9] F. A. Horak, J. Appl. Phys. **23**, 346 (1952).
[10] F. A. Horak, U. S. Office of Naval Research Tech. Report No. 7, University of Missouri (1951).

TABLE I. Solid-state reactions between interface compounds and metallic impurities.

	W	Ti	Si	Al
Ba_3WO_6	0	+	+	+
Ba_2TiO_4	−	0	+	+
Ba_2SiO_4	−	−	0	+
$BaAl_2O_4$	−	−	−	0

interface failed to reveal Ba_3WO_6 which was present early in life, but in its place was found Ba_2SiO_4. A chemical analysis of the original base alloy indicated the presence of 0.02 percent silicon in addition to 4.7 percent tungsten. Thus it is postulated that although the Ba_3WO_6 had been formed initially,[11] it had in time reacted with the silicon to yield Ba_2SiO_4. This possibility is of interest since the reaction with silicon releases tungsten which may again react with BaO producing excess barium. If this does not occur, the displaced tungsten itself may activate the cathode coating and interface. This process is evidently one that proceeds slowly during the life of the cathode and may be responsible for the superior electron emission properties of these cathodes.

EXPERIMENTAL

In order to determine the validity of the assumed silicon–Ba_3WO_6 reaction as well as others of this type, a series of solid state chemical reactions were tested. The four known interface compounds were synthesized[12] and then heated with the metallic impurities tungsten, titanium, silicon, and aluminum to determine if the reaction occurs under conditions similar to those found during normal cathode operation. Each metallic impurity in the form of a fine powder less than 325 mesh in size was ground thoroughly with one of the interface compounds. These two components were used in such proportions as to yield the suspected interface compound. This mixture, usually about 500 mg, was placed in a platinum lined tantalum crucible and heated in vacuum for two to three hours at 1050°C. The pressure was maintained at a value less than 10^{-4} mm of mercury at all times and was usually 10^{-6} mm at the end of the reaction period. X-ray diffraction analysis techniques were used to identify the reaction products. Table I shows the results obtained from this series of twelve reactions. The + sign indicates that the reaction did take place and the − sign indicates that no reaction was observed. For example, titanium, silicon, and aluminum react with Ba_3WO_6 to yield Ba_2TiO_4, Ba_2SiO_4, and $BaAl_2O_4$, respectively. Similarly, silicon and aluminum will replace titanium in the interface compound Ba_2TiO_4, but only aluminum will replace silicon in Ba_2SiO_4. Thus the interface compounds of Table I are listed in order of chemical stability with Ba_3WO_6 as the least stable.

All reactions in which aluminum enters one of the interface compounds yield free barium, for example,

$$2Al+Ba_2SiO_4 \rightarrow BaAl_2O_4+Si+Ba.$$

The production of barium was evident on the glass walls of the vacuum chamber in each case, and the loss of weight of the reacting sample confirmed this production. For those reactions in which tungsten, titanium, or silicon were released, the x-ray diffraction patterns confirmed their presence. One reaction between silicon and Ba_3WO_6 was carried out at 900°C for two hours. The reaction product contained silicon, tungsten, Ba_3WO_6, and Ba_2SiO_4, indicating that the reaction had not reached completion at this lower temperature.

CONCLUSION

The results shown in Table I confirm the original assumption that silicon present in the cathode base metal may replace tungsten in the interface compound. As a result of the relatively high concentration of tungsten as compared to silicon in the original base metal, the tungsten interface was formed first and was then slowly converted to the silicon interface as the silicon diffuses to the interface region. At the completion of the 6000 hour life test these cathodes showed no interface resistance, whereas similar cathodes which also have the Ba_2SiO_4 interface but were formed on a silicon-nickel base exhibited a high resistivity. However, x-ray patterns show no great difference in interface thickness of these two samples at 6000 hours. Thus one concludes that the silicate interface formed from the tungstate interface is in a higher state of activity than is the silicate interface formed by the direct combination of silicon and the oxide coating. The interface activation mechanism, be it due either to excess barium or excess tungsten, may also be responsible for the high cathode-coating activity and its long life.

The results of this study point to two previously neglected factors in oxide cathode technology: (1) small concentrations of certain minor impurities in the base metal may, under conditions of extended life test, ultimately control both the nature of the interface and the thermionic activity of the coating; (2) the judicious selection of two or more base metal impurities in the optimum concentrations may result in higher activity, longer life cathodes than are presently being made.

[11] This interface was probably formed during the conversion of the cathode by the reaction $3BaCO_3+W \rightarrow Ba_3WO_6+3CO$: see Hughes, Coppola, and Evans, J. Appl. Phys. **23**, 635 (1952).

[12] These compounds were produced by reacting $BaCO_3$ with either SiO_2, WO_3, TiO_2, and Al_2O_3 in air at 1300°C.

JOURNAL OF APPLIED PHYSICS VOLUME 24, NUMBER 5 MAY, 1953

Spectral Emissivities of Graphite and Carbon

R. J. THORN AND O. C. SIMPSON
Chemistry Division, Argonne National Laboratory, Lemont, Illinois
(Received December 31, 1952)

The spectral emissivities of graphite and carbon have been determined at a wavelength of 0.653μ in the temperature range of 1285 to 2035°K and have been found to be temperature dependent in accordance with a prediction from a relation between reflectivity and resistivity. For polished graphite of high purity having a resistivity of about 1120 microhm-cm the value obtained is 0.78; the value obtained for a polished high density graphite having a resistivity of 1740 microhm-cm is 0.78; and that for polished spectroscopic carbon having a resistivity of 5080 microhm-cm is 0.79. If graphite is sublimed from a polished surface, the latter becomes roughened, so that the emissivity increases. In one particular experiment it attained a value of 0.90.

I. INTRODUCTION

SINCE an optical pyrometer is frequently used to measure the temperature of the surface of graphite, one of the most important high temperature properties of graphite is its spectral emissivity. The measurements given in the literature[1–3] either have not been made with the proper regard to the variation among graphites or are not sufficiently complete from the point of view of surface conditions to permit precise temperature measurements of many graphites currently used. Quite aside from these criticisms there is in addition a question concerning the temperature coefficient of emissivity. Mendenhall and Forsythe's data on *carbon* together with Chaney, Hamister, and Glass's value appear to indicate a temperature independency for carbon in the range of 1400 to 2400°C, whereas Prescott and Hincke obtained for *graphite* a value of -5.8×10^{-5} for de_λ/dT. One might well inquire whether this difference is real. One of the most common errors encountered in emissivity measurements is that resulting from the presence of temperature gradients, and since this error is frequently such as to cause the measured emissivity to decrease with increasing temperature, one becomes suspicious of a negative coefficient.

The suggestion that Wien's law ($J_\lambda=\text{const}\lambda^{-5}e^{-c_2/\lambda T}=e_\lambda c_1\lambda^{-5}e^{-c_2/\lambda T}$) describes the behavior of the emittance of radiant energy from a nonblackbody with respect to its temperature dependence carries with it the assumption that the spectral emissivity is independent of temperature. Hence a determination of the emissivity as a function of temperature is a method of examining the validity of Wien's law for nonblackbodies. Furthermore, it is desirable to have experimental data to compare with the rough values of the temperature coefficients estimated from electromagnetic theory. The Drude relation between reflectivity and resistivity is $1-e_\lambda=R_\lambda\cong1-2(\rho\nu)^{\frac{1}{2}}$, where ρ is the resistivity and ν is the optical frequency. From this one finds that $(de_\lambda/dT)\cong(\nu/\rho)^{\frac{1}{2}}(d\rho/dT)$. For most artificially prepared graphites $d\rho/dT$ is positive above about 400°C; hence de_λ/dT should be *positive*. Using the more exact theoretical relation between reflectivity and resistivity[4] one finds that for artificial graphite the calculated value is $de_\lambda/dT=2\times10^{-5}$, whereas for carbon the value is -6×10^{-7}. Admittedly the Drude relation may not be well obeyed for wavelengths shorter than 25μ. However, there is no indication of color or anomalous dispersion in the visible region of the spectrum in thin evaporated layers of graphite. Hence the Drude expression may not be too bad an approximation even for visible wavelengths.

II. THEORETICAL BACKGROUND

A. Method

In principle, spectral emissivities are readily obtained from the relation

$$\frac{1}{T_s}-\frac{1}{S_b}=\frac{\lambda}{c_2}\ln e_\lambda, \tag{1}$$

in which the brightness temperature S_b of a surface at a true temperature T_s is measured easily with the aid of an optical pyrometer. However, primarily because of temperature gradients, the determination of the exact true temperature of the surface T_s becomes a difficult problem. The method used to obtain the true surface temperatures given in this report is in principle the same as that employed by Mendenhall and Forsythe, and Prescott and Hincke. A graphite cylinder was heated electrically, and the true surface temperature was obtained by correcting the blackbody temperature of the inside of the cylinder for gradients through the wall. The method of obtaining the correction, however, differs at least in detail from those employed in the two previous investigations. To effect this correction Mendenhall and Forsythe measured the brightness temperature and the corresponding hohlraum temperature as a function of wall thickness and extrapolated the curve of the temperature differences to zero wall. Prescott and Hincke calculated the correction for finite wall thickness from data related to the thermal con-

[1] C. E. Mendenhall and W. E. Forsythe, Astrophys. J. **37**, 380 (1913).
[2] C. H. Prescott, Jr., and W. B. Hincke, Phys. Rev. **31**, 130 (1928).
[3] Chaney, Hamister, and Glass, Trans. Electrochem. Soc. **67**, 138 (1935).
[4] M. Abraham and R. Becker, *Electricity and Magnetism* (Blackie and Son, Ltd., London, 1937), p. 192.

ductivity which they obtained during their investigation. In the authors' investigation the correction was found by extrapolating the temperature measurements *linearly* with the aid of a *theoretically* derived expression in which the temperature difference is not linear with wall thickness. This method is superior to those previously used in that a correct linear extrapolation is used, and no knowledge of the thermal conductivity is needed. Besides this difference, the study described herein also gives more reliable results in that the existence of a "hot" zone surrounding the blackbody hole is recognized and is also taken into consideration in the calculations. Because of this "hot" zone, the area adjacent to the hole cannot be used to measure the brightness temperature of the surface. Since Prescott and Hincke measured the surface temperature between two holes which were rather close together, their emissivities are probably in error, particularly at the higher temperatures. Employment of measurements (without corrections) adjacent to the hole, together with experimental values for the thermal conductivity, can result in calculated emissivities which are greater than unity. Quite possibly this is the difficulty encountered by Prescott and Hincke which forced them to conclude that employment of accepted values for the thermal conductivity gave absurd results.

To apply the corrections for the temperature gradients, obviously one should employ a set-up which approaches as nearly as possible an ideal one. For this reason the inductive rather than the resistive method of heating the graphite was employed. Its use facilitated construction of the apparatus so that the temperature gradient in the radial direction could be minimized and was eliminated entirely in the axial direction, since the sample was mounted so that conduction losses were negligible.

B. Quantitative Relations for Idealized Systems

The idealized system, then, which one wishes to describe quantitatively is one consisting of a hollow graphite cylinder of infinite length (from the point of view of heat losses out the ends) placed in a vacuum and mounted coaxially with an induction coil. This coil is assumed to be long compared with its own radius and with the length of the cylinder and is assumed to carry a sinusoidally varying current. For such a system the heat flow in the graphite cylinder is entirely radial so that the heat flow equation becomes

$$\frac{1}{r}\frac{d}{dr}\left(r\frac{dT}{dr}\right)=-\frac{P(r)}{K}, \tag{2}$$

in which $P(r)$ is the heat generated per unit volume at radius r averaged over each cycle of the sinusoidal induction current, and K is the thermal conductivity. To solve Eq. (2) one must first obtain an expression for the time averaged heat generated as a function of the magnetic intensity arising from the current in the coil. This expression can be found readily to be

$$P(r)=\frac{\rho}{16\pi^2}\left|\frac{dH}{dr}\right|^2, \tag{3}$$

in which ρ is the electrical resistivity in electromagnetic units and H is the root mean square of the real and imaginary amplitudes of a sinusoidally varying complex magnetic intensity.*

The radial dependence of H can be found by solving the equation

$$\frac{d^2H}{dr^2}+\frac{1}{r}\frac{dH}{dr}-\frac{i4\pi\mu\omega}{\rho}H=0, \tag{4}$$

in which μ is the magnetic permeability, and ω is the angular frequency.

To obtain a quantitative expression for the temperature gradient across the wall of the cylinder one first solves Eq. (4) in terms of Bessel functions,[5] substitutes the solution with the aid of Eq. (3) into Eq. (2), and then solves this equation. The result obtained is some function involving the outside and inside radii a and b respectively, the total emissivity, the permeability, the frequency, the thermal conductivity, and the electrical resistivity as parameters.[6] Hence the solution to Eq. (2) can be written in the form

$$T_b-T_a=\frac{C(e,\mu,\rho,\omega)}{K}T_a^4 f(ma,mb), \tag{5}$$

in which T_b is the temperature inside the cylinder,

$$C=\frac{\sigma\epsilon}{m},$$

$$m^2=\frac{4\pi\mu\omega}{\rho},$$

$$f(ma,mb)=ma\left\{\frac{\int_{mb}^{ma}\frac{1}{z}\int_{mb}^{z}z\left|\frac{dH}{dz}\right|^2 dz\,dz}{\int_{mb}^{ma}z\left|\frac{dH}{dz}\right|^2 dz}\right\}, \tag{6}$$

and $z=mr$.

The radial function $f(ma,mb)$ as defined by Eq. (6) is a complicated expression involving Bessel functions of

* That is, if the magnetic intensity in phase and amplitude is given by $\bar{H}=(H_1+iH_2)e^{i\omega t}|_{\text{real}}$, then the proper H to put into Eq. (4) is $H=(H_1^2+H_2^2/2)^{\frac{1}{2}}$. Because of the symmetry of the problem, H is a function of the radial coordinate only.

[5] N. W. McLachlan and A. A. Meyers, Phil. Mag. **19**, 846 (1935).

[6] R. J. Thorn and O. C. Simpson, J. Appl. Phys. **24**, 297 (1953); also ANL-4878, August, 1952, contains mathematical details and other details of the emissivity investigation not included in the present publication.

first and second kind of zero order.[6] Approximately it is given by the expression

$$f(ma, mb) \cong \frac{ma\left[a^4 - b^4\left(1+4\ln\frac{a}{b}\right)\right]}{4(a^4-b^4)}.$$

Equation (5) must be combined with Eq. (1) so that an expression involving T_h and S_b is obtained. To accomplish this one solves Eq. (5) for $1/T_s$ to find

$$\frac{1}{T_s} = \frac{1}{T_h}\left\{1+\frac{CT_s^4}{KT_h}f(ma, mb) + \left[\frac{CT_s^4}{KT_h}f(ma, mb)\right]^2 + \cdots\right\}. \quad (7)$$

Since $[CT_s^4 f(ma, mb)/KT_h]$ is of the order of 10^{-3} or less, one can neglect all such terms with degree higher than one. If this value of $1/T_s$ given by Eq. (7) is substituted into Eq. (1), one obtains

$$\frac{1}{T_h} - \frac{1}{S_b} = \left(\frac{\lambda}{c_2}\right)\ln\epsilon_\lambda - \frac{CT_s^4}{KT_h^2}f(ma, mb). \quad (8)$$

The left side of this expression involves the temperatures T_h and S_b both of which are readily measured with an optical pyrometer. At constant hohlraum temperature T_h and constant outside radius a and variable inside radius b the ratio of T_s^4/T_h^2 will vary only slightly as $f(ma, mb)$ varies. Hence, according to this expression, the difference between the reciprocal temperatures $(1/T_h - 1/S_b)$ is a linear function of $f(ma, mb)$ with an intercept at $f(ma, ma)=0$ equal to $(\lambda/c_2)\ln\epsilon_\lambda$ and a slope inversely proportional to the thermal conductivity. By calculating the intercept one can then obtain the spectral emissivity without having a knowledge of the thermal conductivity. On the other hand, the slope gives knowledge of the total emissivity and the thermal conductivity.

C. Hot Zone Surrounding Blackbody Hole

Since the induced current lines are crowded together around the blackbody hole in the side of the cylinder, the area adjacent to the hole may be slightly hotter than the rest of the cylinder. (This effect is even more pronounced for the area between two closely located holes.) For this reason an area symmetrically placed with respect to the hole, but 180° from it ("backside") was used to determine the brightness temperature of the surface. However, as shown below, this crowding of current may also cause the temperature of the hole as measured by the pyrometer to be slightly higher than the hohlraum temperature. If one for simplicity considers the multiple reflection of a ray of light between two parallel planes, at the same temperature T_h except for an area adjacent to the hole in one of the planes which is assumed to be at a higher temperature T_f, then one can calculate the intensity of radiation emerging from the hole. The normal spectral radiancy of a ray of light originating at an internal point A some distance from the hot zone is $e_\lambda J(T_h)$, where $J(T_h)$ is given by Wien's law, and T_h is the true hohlraum temperature of the space between the plates. After reflection at some internal point B, the radiancy is $e_\lambda J(T_h)+R_\lambda e_\lambda J(T_h)$; after reflection at internal point C it is $e_\lambda J(T_h)+R_\lambda e_\lambda J(T_h)+R_\lambda^2 e_\lambda J(T_h)$. Continuing in this manner one finds that the radiancy after reflection at internal point Y adjacent to the hole and in the hot zone is $e_\lambda J(T_f)+R_\lambda e_\lambda J(T_h)+R_\lambda^2 e_\lambda J(T_h)+R_\lambda^3 e_\lambda J(T_h)+\cdots$, and after Z opposite the hole is

$$J_p = e_\lambda J(T_h)+R_\lambda e_\lambda J(T_f) + R_\lambda^2 e_\lambda J(T_h)+R_\lambda^3 e_\lambda J(T_h)+\cdots,$$

where J_p is the spectral radiancy of the hole as seen by the pyrometer. This is proportional to the intensity of light entering the pyrometer. Since $e_\lambda = 1-R_\lambda$, one finds that

$$J(T_h) = \frac{J_p - \{R_\lambda - R_\lambda^2\}J(T_f)}{\{1-R_\lambda+R_\lambda^2\}}. \quad (9)$$

In deriving Eq. (9) we have not considered in detail the actual area of the "hot" zone. One can readily show, however, that this will cause no large error. If one considers the case of two parallel plates, one with the hole at a temperature T_f and the other at a temperature T_h, then the expression analogous to Eq. (9) is

$$J(T_h) = \frac{J_p - \{R_\lambda - R_\lambda^2 + R_\lambda^3 - \cdots\}J(T_f)}{\{1-R_\lambda+R_\lambda^2-R_\lambda^3+\cdots\}}. \quad (10)$$

If $R_\lambda = 0.2$, $T_h = 1905°\text{K}$ and $T_f = 1920°\text{K}$, then Eq. (9) gives $T_p = 1907.5°\text{K}$, whereas Eq. (10) gives $T_p = 1907.6°\text{K}$. Therefore the difference is sufficiently small so that it can be neglected, and either Eq. (9) or (10) can be used to correct the pyrometer reading for the "hot" zone so that the true temperature of the hohlraum can be obtained.

III. APPARATUS AND PROCEDURE

A cross sectional drawing of the essential parts of the apparatus is shown in Fig. 1. The arrangement was such that it permitted the pyrometer to be pivoted about the axis of the graphite or carbon cylinder under investigation so that the temperature of the hole and an area diametrically opposite the hole could be determined with one pyrometer. Also the cylinder could be rotated about its axis with the aid of the ground glass joint. The graphite radiation shield could be moved up and down by moving the counterbalance weights with a set of magnetics.

Cylinders were fabricated from a large block of high purity, medium density graphite, from a small sample of a very high density graphite obtained from Great Lakes

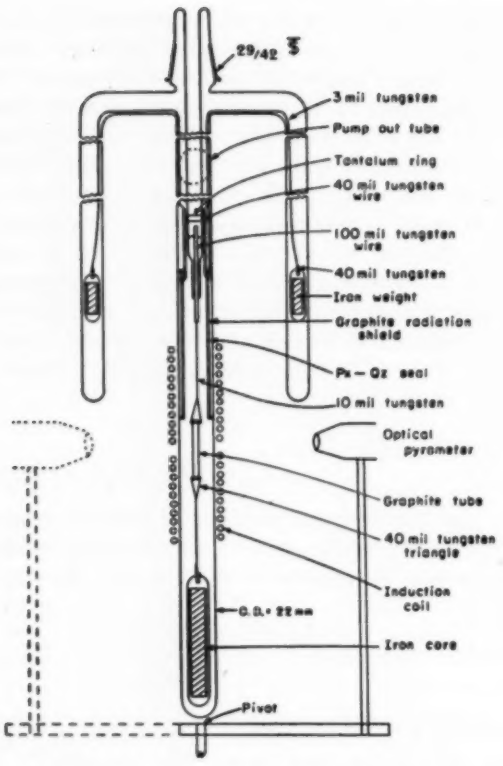


FIG. 1. Apparatus used to determine emissivity.

Carbon Corporation, and from special spectroscopic carbon electrodes obtained from National Carbon Company. The cylinders were hollow with an outside diameter of about 0.5 cm and an initial inside diameter of 0.2578 cm. They were 4 cm long and had a small radial hole, 0.046 cm (No. 77 drill) in diameter, through the wall at the center of the axial length. The axis of the cylinder was cut parallel with the axis of extrusion. When the graphite cylinder axis was cut perpendicular to the extrusion direction, the polishing operation produced a tube which was out-of-round because of the difference in hardness of the graphite in the two directions. The diameter was selected mainly as a compromise between two opposing factors. On the one hand, it was desired to minimize the difference in temperature between the inside and the outside of the cylinder; on the other hand it was desired that the graphite cylinder be sufficiently large so that it would be heated to about 2400°K with the induction heating unit at the authors' disposal. The selection of the length was somewhat arbitrary, but emissivity measurements *vs* cylinder length demonstrated that the measured emissivity was independent of the length over a range of 2 to 12 cm.

The wall thickness of the cylinder was decreased by increasing (by drilling) the inside diameter. To accomplish this without breaking the cylinder it was necessary to use drills the ends of which were ground flat so that there was no tendency for the drill to slip into the cylinder and bind. For the drilling operation the cylinder was held in a brass sleeve which in turn was clamped in a collet in a lathe. The inside diameter was drilled successively with number 38, 31, 29, 23, and 18 drills.

In the investigation being reported two different surface conditions were studied. These were a polished surface and a polished surface from which graphite had been sublimed. The polishing was accomplished by holding wet ground-glass plates against the cylinder (one on each side) as it turned in a lathe. Three sets of plates were used. One was ground with 120 Norbide, another with 600 Carborundum, and the third with 95 Aloxite (used with no water). The surface obtained was one with a black metallic luster (reflectivity about 0.22) perforated with a few small holes which originally were the pores of the piece. This surface was carefully protected during all subsequent operations so as to avoid producing scratches or other surface irregularities.

All cylinders were given a preliminary bakeout at about 1800°K for approximately three hours before any measurements were made. During this treatment a small amount of black material sublimed out of the graphite. After this, no sublimation of material from the cylinder was observed, even at temperatures as high as 2023°K. At all times when the cylinders were heated the pressure was never permitted to exceed 10^{-4} mm Hg. During the temperature measurements the pressure was 5×10^{-6} mm Hg or less. Under these conditions there was no evidence indicating a roughening of the surface by gases in the vacuum.

The optical pyrometer (Leeds and Northrup disappearing filament type) was standardized against a tungsten ribbon lamp calibrated by the National Bureau of Standards. Although the maximum uncertainty in the lamp calibration with respect to absolute temperature may have been as large as 5°, this large an error in the absolute value of the temperature causes an error of only 0.001 in the measured emissivity and hence is not serious. The important error is the relative error in the temperatures T_s and T_A. This is estimated to be about 0.67°. This estimate is based upon the assumption that the temperature of the ribbon in the lamp is a monatonic function of the current through it.† If this is true, then the error for a small range in temperature in the standardization experiment is, for all practical purposes, the error involved in comparing the pyrometer with the lamp. This error causes an error of 0.006 in a single determination of the emissivity at 1600°K.

Since the pyrometer was calibrated against the National Bureau of Standards pyrometer (indirectly), the corrected temperatures correspond to the effective wavelengths of the latter pyrometer. Hence the emissivities reported in this investigation also correspond to

† For example, it is assumed that there are no sudden changes in emissivity, transmissivity, or electrical resistivity due to sudden evaporations or condensations of layers of adsorbed gases within the standard ribbon filament lamp envelope (on the ribbon or on the glass) which cause non-monatonic temperature changes with current.

these same wavelengths. For the mid-point temperature this wavelength is $0.652_6\mu$ and its variation over the range of measurements is $0.003_6\mu$.

The optical transmission of the quartz tube envelope through which the heated cylinder was viewed with the pyrometer was determined by measuring the brightness temperature with the pyrometer of a 6-watt General Electric fluorescent lamp viewed directly and with the quartz tube interposed. This transmission was used to correct the pyrometer readings to find the true temperature of the inside of the cylinder.

This procedure, however, did not permit a sufficiently accurate measurement of the difference between the corrections for the two sides of the quartz tube. These are the corrections which are involved in the calculation of the emissivity from the measured temperatures. If the measurements of the temperatures of the hole and the back side of the cylinder had been made leaving the pyrometer in a fixed position and rotating the cylinder by means of the ground joint, no correction would have been necessary. This procedure, however, would have been time-consuming and perhaps impossible because the joint usually stuck after a few turns. Hence, this difference was determined by measuring the temperature of the hole in its original position with respect to the quartz tube and the temperature of the hole after rotation through 180 degrees; that is, the temperature of the hole measured through the back side of the quartz tube.

IV. RESULTS

A. Temperature Dependence

1. Graphite

Typical values of $1/T_\lambda - 1/S_b$ are plotted with variable inside radius b in Fig. 2 as a function of $f(ma, mb)$ for three different temperatures. Within the scatter of the data these curves may be considered to be linear as demanded by Eq. (8). From the intercepts the emissivities were calculated. The values so obtained are plotted in Fig. 3, together with previously published

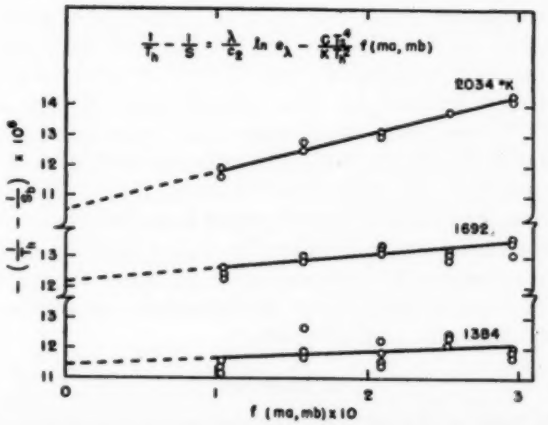


FIG. 2. Plot of reciprocal temperatures *vs* radial function.

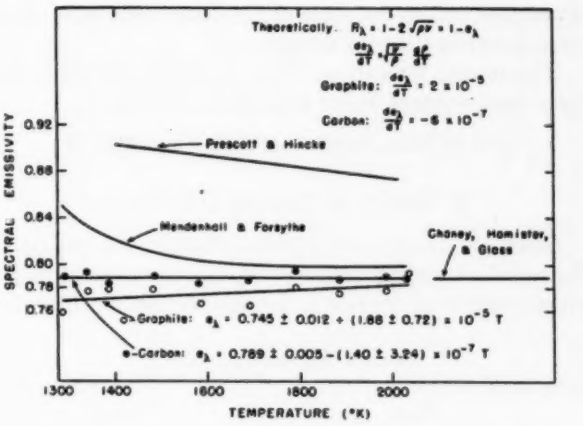

Theoretically
Graphite:
Carbon:
Prescott & Hincke
Mendenhall & Forsythe
Cheney, Homister, & Gloss
Graphite: ϵ_λ = 0.745 ± 0.012 + (1.88 ± 0.72) × 10^{-5} T
Carbon: ϵ_λ = 0.789 ± 0.005 − (1.40 ± 3.24) × 10^{-7} T
SPECTRAL EMISSIVITY
TEMPERATURE (°K)

FIG. 3. Spectral emissivity *vs* temperature.

values. Inspection of these presently reported values shows no trend with temperature other than possibly a linear one. Hence a least-squares linear curve has been calculated. The resulting equation is (for polished graphite) $\epsilon_\lambda = (0.745 \pm 0.012) + (1.88 \pm 0.72) \times 10^{-5} T$, (for $\lambda = 0.653\mu$).

2. Carbon

Because of the difference between the properties of carbon and those of graphite, it was impossible to use the same procedure for carbon that was used for graphite. It was impossible for example to enlarge the inside diameter of a carbon cylinder beyond about a No. 29 drill by the same procedure used with graphite because the cylinder split. Therefore it was decided that measurements with a cylinder having an inside diameter of 0.4296 cm (No. 18 drill) and an outside diameter of 0.5024 cm‡ could be made and then corrected for the slight error caused by the temperature gradient across the wall on the basis of the data obtained with graphite. From the thermal conductivities and the resistivities of the two materials it was found that the slope of $1/T_\lambda - 1/S_b$ *vs* $f(ma, mb)$ as given by Eq. (8) is about 1.21 times larger for carbon than it is for graphite. Hence the correction at a given temperature for carbon for the extrapolation from $f(1.49, 1.28)$ to $f(1.49, 1.49)$, is about 1.21 times that for graphite. In addition to this difference it was also found that with the apparatus used for graphite it was impossible to heat the carbon cylinder above about 1500°K. Hence to obtain measurements at higher temperatures it was necessary to surround the carbon cylinder with a thin-walled graphite cylinder, snugly fitted, with holes (No. 60 drill) in the walls so that the blackbody hole and the backside of the carbon cylinder could be seen. To ascertain whether this encasing cylinder had any effect on the measured emissivity, measurements were made up to 1500°K before and after the graphite was placed around the carbon. The result with carbon alone is 0.794±0.005, whereas that with the combination is 0.788±0.005.

‡ A cylinder with this thin wall could be made by first drilling the inside out and then turning the outside down on a lathe.

Therefore within the limit of the error there is no difference between the two set-ups.

The results for carbon (Fig. 3) can be expressed also by a least-squares linear equation—namely,

$$e_\lambda = (0.789 \pm 0.005) - (1.40 \pm 3.24) \times 10^{-7} T.$$

B. Effect of Density and Polish

Since one might well inquire whether the emissivity for polished graphite reported in the previous section is characteristic of the particular sample used or the extent of the polishing, data pertaining to these effects are reported below. In one series of experiments the values of T_λ and S_b were determined at 1480°K for a graphite cylinder having the same length and outside radius as that used for the previously reported experiment, but having an inside radius of a No. 32 drill. Also this new sample was fabricated from a piece of high density artificial graphite obtained from Great Lakes Carbon Corporation. To ascertain the correction for the $f(ma, mb)$ extrapolation, the values of $(1/T_\lambda - 1/S_b)$ were measured at two temperatures (1480 and 2030°K) in one particular case. From the difference between these values and the corresponding difference obtained with the medium density graphite, the ratio of the two corrections for the $f(ma, mb)$ extrapolation was obtained. The correction in this particular case was then obtained from the one previously determined. Since the correction for the hot zone surrounding the hole was only about 0.003 in the first experiment, it was assumed to be the same in this experiment. The emissivity of the high density graphite polished with plates ground with 95 Aloxite was determined; the cylinder was removed from the apparatus, repolished with 95 Aloxite plates, and the emissivity redetermined. Finally, the cylinder was polished with plates ground with sapphire powder (0.5–3.0μ particle size) and its emissivity remeasured. The results (Table I) show that further polishing, even with the sapphire powder plates, beyond the 95 Aloxite plates has no effect on the emissivity. Hence, for the high density graphite, one obtains for the emissivity a value of 0.780 ± 0.005.

TABLE I. Emissivity vs polish.

Polish	Emissivity
95 Aloxite plates	0.777 ± 0.009
Repeat polish	0.784 ± 0.009
Sapphire powder plates	0.780 ± 0.009

TABLE II. Emissivity vs resistivity for "carbons."

Material	Resistivity (microhm-cm)	Emissivity
a axis graphite	36[b]	*ca* 0.69[b]
Medium density ⊥[a] (*ca* 1.6 g/cm³)	1120[c]	0.78
Great Lakes high density ⊥[a] (*ca* 1.9 g/cm³)	1740	0.78
Spectrochemical carbon	5080[d]	0.79
c axis graphite	2 000 000[e]	*ca* 0.95 (?)[f]
	5000[b]	

[a] ⊥ = viewed perpendicular to extrusion axis.
[b] L. Fuchs, private communication.
[c] Neubert, Royal, and VanDyken, Argonne National Laboratory Report, ANL-4022 (August, 1947).
[d] H. S. Patten, private communication.
[e] K. S. Krishnan and N. Ganguli, Nature 144, 667 (1939).
[f] Schneiderhohn and Ramdohr, *Lehrbuch der Erzmikroskopie* (Bornhaeger, Berlin, 1931), Vol. 2, p. 18.

This value is very slightly higher than that obtained for the medium density. In view of the fact that the electrical resistance of the high density graphite is higher than that of the previous graphite sample, and also in view of the fact that the area observed on the latter was carefully selected to be relatively free of holes, it is believed that the difference is real. This conclusion is supported by the fact that this value falls between that for graphite and that for carbon. Hence, as one can predict from the Drude relation, the emissivity of the various "carbons" increases with increasing resistivity. It is interesting to observe that even the values for single crystals of graphite also fulfill this relation qualitatively (see Table II).

C. Effect of Sublimation

Since one of the primary factors which encouraged this investigation was the need for the accurate determination of the temperature in the measurement of the heat of sublimation of graphite, it was necessary to find the effect of sublimation from some reproducibly prepared surface on the emissivity of that surface. Consequently, the emissivity of a polished surface (95 Aloxite plates) of a cylinder (No. 38 drill inside diameter; 0.5009 cm outside diameter; 4 cm long) fabricated from the medium density graphite was determined first before sublimation and then after a number of sublimations at a background pressure of 10^{-5} mm of Hg. To attain the sublimation temperatures readily and safely, it was necessary to use a thick walled cylinder, the radiation shield, and water cooling of the quartz tube. The latter was accomplished by surrounding the quartz and induction coil with a large Pyrex tube through which distilled water was circulated by means of a small pump. The values of T_λ and S_b were determined after each sublimation at two temperatures (1480 and 2030°K) and each set of values was corrected for the extrapolation to zero wall thickness and for the hot zone according to the procedure outlined in the previous section. The emissivities calculated from these corrected temperatures are plotted in Fig. 4 as a function of the total amount of graphite sublimed. Although the scatter of the points is rather large, the emissivity approaches a fairly constant value of 0.905.

V. ERRORS

In view of the fact that the measurements have been extrapolated to cylinders having zero wall thickness and

that radiation inside the cylinder is very nearly that of a blackbody, the errors in the emissivity are easily shown to arise primarily from the precision of the measurements with the pyrometer and not because of failure of the method to approach ideal conditions. Hence the scatter of the data measures the error. The least-squares method has therefore been used to calculate their magnitudes as given in the results.

VI. ADDENDUM

A. Thermal Conductivity of Graphite

According to Eq. (8) the slope of $1/T_h - 1/S_b$ *vs* $f(ma, mb)$ is inversely proportional to the thermal conductivity. Since C in Eq. (8) is equal to $\sigma\epsilon/m$, and since T_s^2/T_h^2 is close to unity, one obtains for the thermal conductivity

$$K = \sigma\epsilon T_s^2/\text{cm},$$

in which c is the slope of the $1/T_h - 1/S_b$ *vs* $f(ma, mb)$ curve. If the values of the slope are substituted into the equation given above together with the value of $\epsilon = 0.80$, $\rho = 8.39\times10^5$ emu, and $\omega = 2\pi\times3.75\times10^5$ sec^{-1}, one obtains an average value of 0.16 cal cm^{-1} sec^{-1} deg^{-1} for the thermal conductivity. Although the individual values scatter quite a great deal, they are at least of the correct order of magnitude. For example, Powell[7] has reported a value of about 0.05 for this temperature range and MacPherson[8] has reported a value of about 0.11 for AGW graphite. In view of the small temperature differences involved in the calculation no great accuracy can be expected. However, the values obtained do indicate that one is justified in employing Eq. (8) to calculate the emissivity. Perhaps if the above method were improved by employing cylinders with thicker walls it could be used to measure the thermal conductivity of graphite at high temperatures.

C. Correction to Heat of Sublimation

The value for the emissivity of a surface from which graphite had been sublimed as determined in the present study will have no effect on the value for the heat of sublimation obtained by Simpson, Thorn, and Winslow[9]—namely, 171.5±1.9 kcal/mole (this is the average of 174±3 and 169±2 weighted equally), since temperatures of cylindrical holes were measured to obtain this value. However, the value obtained herein for e_λ will change the value of the heat of sublimation of 175.2 kcal/mole reported by Marshall and Norton.[10] The average temperature of their measurements giving this average ΔH_0^0 is 2569°K. If one uses a value of 0.905 typical for the emissivity of graphite under conditions of sublimation instead of the value 0.80 used by Marshall and Norton, and if one corrects their temperature to the 1948 scale,[11] then one finds that

$$\Delta H_0^0 + 3.0 = \frac{178.2\times2520}{2569},$$

$$\Delta H_0^0 = 171.8\pm0.24 \text{ kcal/mole}.$$

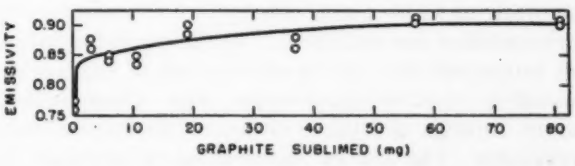


FIG. 4. Emissivity as a function of amount sublimed.

Hence there is excellent agreement between these two determinations for the heat of sublimation, and both agree reasonably well with the value of 170.6±0.2 from spectroscopic data favored by Gaydon[12] but not with the corrected value of 165.3 kcal/mole reported by Brewer and Mastick.[13]

The authors wish to acknowledge the helpful suggestions given to them by B. I. Spinrad and G. H. Winslow.

[7] R. W. Powell, Proc. Phys. Soc. (London) **49**, 419 (1937).
[8] D. Cowen, NEPA-953, p. 321.
[9] Simpson, Thorn, and Winslow, Argonne National Laboratory Report ANL-4264 (May, 1949).
[10] A. L. Marshall and F. J. Norton, J. Am. Chem. Soc. **72**, 2166 (1950).
[11] R. J. Corruccini, J. Research Natl. Bur. Standards **43**, 133 (1949).
[12] A. G. Gaydon, *Dissociation Energies and Spectra of Diatomic Molecules* (Chapman and Hall, Ltd., London, 1947).
[13] L. Brewer and D. F. Mastick, University of California Radiation Laboratory Report UCRL-572 (1949).

Axial Motion of Electrons in a Linear Accelerator with $\beta=1$*

JAMES SWIHART AND EDWARD AKELEY
Purdue University, Lafayette, Indiana

(Received January 5, 1953)

Complete integrated solutions have been found in terms of elementary functions for the motion of an electron on the axis of a progressive wave linear accelerator with a constant wave speed, $\beta=1$. The character of the solutions has been discussed and asymptotic relations found.

Calculations have been made for the third section of the Purdue linear accelerator, in which the curves for the first two sections where $\beta<1$ have been extended. The energy distribution of the emergent beam was computed. It was found possible to design a linear accelerator with constant wave speed sections which, without preliminary bunching, would produce a beam of greater homogeneity in energy than might have been expected.

INTRODUCTION

THE motion of an electron on the axis of a progressive wave linear accelerator with a constant wave speed, $\beta=1$, is treated. This is a continuation of the work of Caplan and Akeley[1] who considered the corresponding case with $\beta<1$. The differential equations are integrated and the result applied to calculations extending the energy-distance and phase-distance curves through the third section of the Purdue linear accelerator. The energy distribution is obtained for the entire accelerator from these calculations.

THE EQUATIONS OF MOTION

The differential equations of motion of an electron on the axis of a progressive wave linear accelerator with $\beta=1$, are

$$\frac{du}{dx}=A\cos y$$
$$\frac{dy}{dx}=1-\frac{nu}{(u^2-1)^{\frac{1}{2}}} \tag{1}$$

where u is the rest plus kinetic energy in units of the rest energy of the electron, x is the axial distance in units of c/ω where c is the speed of light and ω is the frequency of the wave, y is the phase given by $(x-t)$ where t is the time in units of $1/\omega$, $n=+1$, -1 for positive and negative velocities, respectively, and A is given by $A=eE/mc\omega$ where e is the charge of the electron, m is the rest mass of the electron, and E is the minimum value of the electric field strength along the axis of the tube (i.e., E is the negative of the maximum value of the field strength so that A is positive since e is negative).

The first integration of this equation is[2]

$$n(u^2-1)^{\frac{1}{2}}-u=C-A\sin y, \tag{2}$$

where C is the constant of integration.

Since from (1)

$$A\sin y=\pm(A^2-A^2\cos^2 y)^{\frac{1}{2}}=\pm\left[A^2-\left(\frac{du}{dx}\right)^2\right]^{\frac{1}{2}}, \tag{3}$$

(2) may be written

$$n(u^2-1)^{\frac{1}{2}}-u=C\mp\left[A^2-\left(\frac{du}{dx}\right)^2\right]^{\frac{1}{2}}, \tag{4}$$

or

$$\frac{du}{dx}=\pm[A^2-(n[u^2-1]^{\frac{1}{2}}-u-C)^2]^{\frac{1}{2}}. \tag{5}$$

The upper sign is for the case of the energy increasing with increase of distance (positive acceleration for positive velocity). The sign corresponding to this case will be carried on top in the following.

Introducing s defined by

$$s=n(u^2-1)^{\frac{1}{2}}-u-C=-A\sin y, \tag{6}$$

and separating variables, we obtain for Eq. (5)

$$dx=\pm\frac{ds}{2(C+s)^2(A^2-s^2)^{\frac{1}{2}}}\mp\frac{ds}{2(A^2-s^2)^{\frac{1}{2}}}. \tag{7}$$

On integrating Eq. (7), we find

$$x=\mp\tfrac{1}{2}\sin^{-1}(s/A)\pm\tfrac{1}{2}B+C_1, \tag{8}$$

where

$$B=\int\frac{ds}{(C+s)^2(A^2-s^2)^{\frac{1}{2}}},$$

and C_1 is the second constant of integration.

B depends on whether the absolute value of C is equal to, greater than, or less than A. The case for $|C|=A$ gives the curve known as the separatrix. Let $w=C+s$; then

$$B=\int\frac{dw}{w^2(-w^2+2Cw+A^2-C^2)^{\frac{1}{2}}}. \tag{9}$$

* This material was presented at the Chicago Meeting of the American Physical Society in 1950 and in Phys. Rev. **81**, 304 (1951). It was also accepted in partial fulfillment of the requirements of a Master's degree at Purdue University in June, 1951. The same subject has also been treated by W. Skolnik, Master's thesis, Illinois Institute of Technology, June, 1950; and by E. L. Chu, M. L. Report No. 140, May, 1951, Microwave Laboratory, Stanford University.

[1] D. Caplan and E. Akeley, J. Appl. Phys. **23**, 774 (1952).

[2] J. C. Slater, Revs. Modern Phys. **20**, 473 (1948).

For $|C|=A$,

$$B=-\int dw w^{-3}(2C/w-1)^{-\frac{1}{2}}. \tag{10}$$

It is shown later that C is negative for this case. The minus sign in front of the integral enters from the fact that $w<0$ since $C<0$ and since $|s|\leq A=|C|$ by Eq. (6). Let $Z^2=2C/w$. Then

$$B=\frac{1}{2C^2}\int\frac{Z^3dZ}{(Z^2-1)^{\frac{1}{2}}}=\frac{1}{2C^2}\left[\frac{(Z^2-1)^{\frac{1}{2}}}{3}+(Z^2-1)^{\frac{1}{2}}\right],$$

or

$$B=\frac{1}{2C^2}\left[\frac{1}{3}\left(\frac{C-s}{C+s}\right)^{\frac{3}{2}}+\left(\frac{C-s}{C+s}\right)^{\frac{1}{2}}\right]. \tag{11}$$

Substituting this into Eq. (8) and also the value $-A$ for C, we find

$$x=\mp\frac{1}{2}\sin^{-1}\frac{s}{A}\pm\frac{1}{12A^2}\left(\frac{A+s}{A-s}\right)^{\frac{3}{2}}\pm\frac{1}{4A^2}\left(\frac{A+s}{A-s}\right)^{\frac{1}{2}}+C_1, \tag{12}$$

for the case of the separatrix.

For the case $|C|>A$

$$B=-\frac{[-w^2+2Cw+(A^2-C^2)]^{\frac{1}{2}}}{(A^2-C^2)w}-\frac{2C}{2(A^2-C^2)(C^2-A^2)^{\frac{1}{2}}}\times\sin^{-1}\left[\frac{2Cw+2A^2-2C^2}{|w|(4C^2+4A^2-4C^2)^{\frac{1}{2}}}\right],$$

or

$$B=\frac{(A^2-s^2)^{\frac{1}{2}}}{(C^2-A^2)(C+s)}+\frac{C}{(C^2-A^2)^{\frac{3}{2}}}\sin^{-1}\left[\frac{A^2+Cs}{A|C+s|}\right]. \tag{13}$$

For the case $|C|<A$

$$B=-\frac{[-w^2+2Cw+(A^2-C^2)]^{\frac{1}{2}}}{w(A^2-C^2)}-\frac{2C}{2(A^2-C^2)}\left[-\frac{1}{(A^2-C^2)^{\frac{1}{2}}}\times\ln\left|\frac{2(A^2-C^2)^{\frac{1}{2}}(-w^2+2Cw+A^2-C^2)^{\frac{1}{2}}}{w}+\frac{2(A^2-C^2)}{w}+2C\right|\right],$$

or

$$B=\frac{(A^2-s^2)^{\frac{1}{2}}}{(C+s)(C^2-A^2)}+\frac{C}{(A^2-C^2)^{\frac{3}{2}}}\times\ln\left|\frac{(A^2-C^2)^{\frac{1}{2}}(A^2-s^2)^{\frac{1}{2}}+(A^2-C^2)+C(C+s)}{(C+s)}\right|. \tag{14}$$

plus a constant times ln2 which can be incorporated in C_1 in Eq. (8). In the above equations, all inverse sine functions are taken from $-\frac{1}{2}\pi$ to $\frac{1}{2}\pi$.

DISCUSSION OF THE EQUATIONS

If p designates the relativistic momentum in units of mc, then

$$p=n|p|=n(u^2-1)^{\frac{1}{2}}. \tag{15}$$

Substituting Eq. (15) into Eq. (2), we find

$$p-(p^2+1)^{\frac{1}{2}}=C-A\sin y. \tag{16}$$

Figure 1 is a plot of Eq. (16) for various values of C. Slater[2] gives a plot of Eq. (2), which has the same type of curves for positive velocities.

Since the speed of the wave is equal to the speed of light and since no particle can attain this speed, the phase y of every particle is always decreasing. Thus the direction of motion in momentum-phase space is as shown by the arrows in Fig. 1.

The constant C can not take all values as is seen by Eq. (16). Since $|p|<(p^2+1)^{\frac{1}{2}}$ for all values of p, $C-A\sin y<0$ and thus $C<A\sin y\leq A$. This means that C can not be positive except for $|C|<A$. Thus the use of $C=-A$ was correct in deriving Eq. (12).

Equation (16) has asymptotes parallel to the momentum axis at the phase values

$$y=\sin^{-1}(C/A)+2\pi k \tag{17a}$$

and

$$y=\pi-\sin^{-1}(C/A)+2\pi k,\quad k=0,\pm1,\pm2,\cdots. \tag{17b}$$

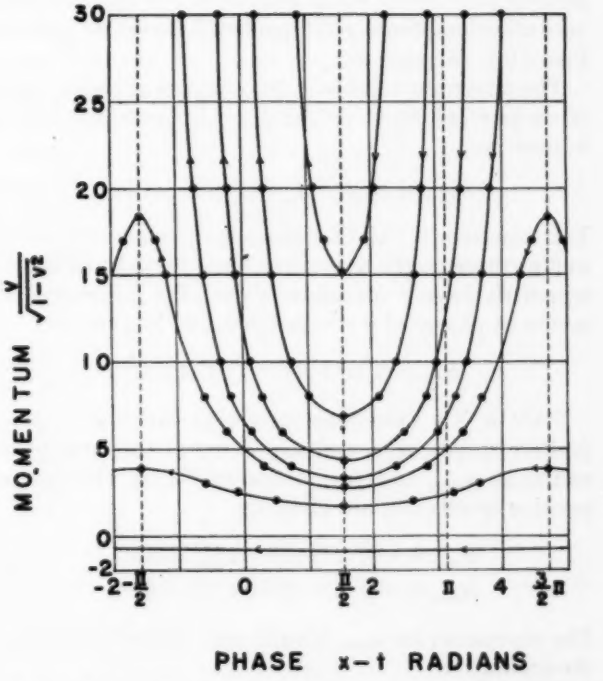


FIG. 1. Momentum of the particle *vs* phase for the case $A=0.0728$. The top curve is for $C=0.0400$, the second for $C=0$, the third for $C=-0.0400$, the fourth for $C=-0.0728$ (the separatrix), the fifth for $C=-0.10$, the sixth for $C=-0.20$, and the bottom curve for $C=-2.00$.

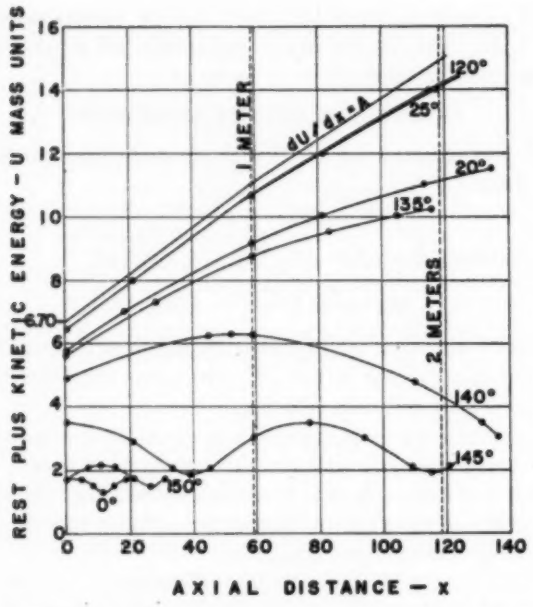


FIG. 2. Rest plus kinetic energy in mass units *vs* axial distance in units of c/ω in the third section for various initial phases in the first section.

The curves of Eq. (16) approach these asymptotes and $p\to+\infty$ only for $|C|\leq A$. Thus Eqs. (12) and (14) are for the cases of $u\to\infty$ as $x\to\infty$. For these cases, y can not take all values but must be in the interval $(\pi-\sin^{-1}\times(C/A)+2\pi k,\ \sin^{-1}(C/A)+2\pi k)$. For the case that $|C|>A$, Eq. (16) shows that p is bounded and is a periodic function of y with period 2π. For this case y can take all values. Since p is bounded, u is also, as given by Eqs. (13), (8), and (6).

The minimum value of p, $p_{\min}$, occurs at phase values of $y=\frac{1}{2}\pi+2\pi k(k=0,\ \pm1,\ \pm2,\ \cdots)$. From Eq. (16), it is given by

$$p_{\min}=\tfrac{1}{2}[(C-A)^2-1](C-A)^{-1}. \tag{18}$$

This equation is valid for particles above and below as well as those on the separatrix. Only particles below the separatrix have a maximum value of p however. This occurs at phases of $y=-\frac{1}{2}\pi+2\pi k$ and is given by

$$p_{\max}=\tfrac{1}{2}[(C+A)^2-1](C+A)^{-1}. \tag{19}$$

Since u is a monotone increasing function of p for positive speeds, $u_{\max}$ and $u_{\min}$ occur at the same phase values as $p_{\max}$ and $p_{\min}$, respectively, for the case of positive speeds and are given by

$$\begin{aligned} u_{\max}&=-\tfrac{1}{2}[(C+A)^2+1](C+A)^{-1},\\ u_{\min}&=-\tfrac{1}{2}[(C-A)^2+1](C-A)^{-1}. \end{aligned} \tag{20}$$

The expression for $u_{\max}$ is valid only for particles below the separatrix.

From Eq. (16) it easily follows that if $C>A-1$, the electron will always have a positive speed. On the other hand if $C<-A-1$, the speed is always negative. Particles for which $-A-1<C<A-1$ will have positive speeds part of the time as well as negative speeds. Thus it is seen that C is an important parameter in determining the type of motion of a given electron.

Equation (8) together with Eqs. (13) and (6) gives the energy as a function of the distance for particles below the separatrix. These equations show that for this case u is a periodic function of x for particles that always have a positive speed and the period is

$$\Delta x=-\pi-C\pi(C^2-A^2)^{-\frac{1}{2}}. \tag{21}$$

Equation (8) together with Eqs. (14) and (6) gives the energy as a function of the distance for particles above the separatrix. Considering values of x so large that we can neglect terms of the order of $1/x$ or higher, we find from these equations

$$u=Gx+H\ln|x|+I \tag{22}$$

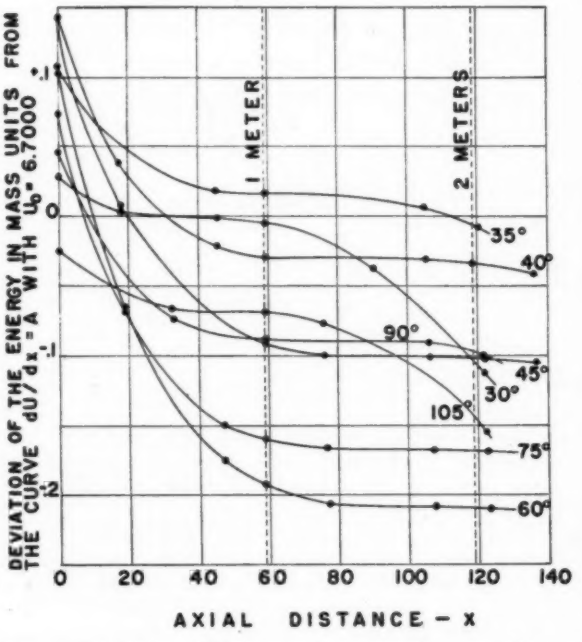


FIG. 3. Deviation of the energy in mass units from the curve $du/dx=A$ with $u_0=6.700$ *vs* axial distance in units of c/ω in the third section for various initial phases in the first section.

where G, H, and I are constants given by

$$\begin{aligned} G&=(A^2-C^2)^{\frac{1}{2}},\\ H&=\tfrac{1}{2}C(C^2-A^2)^{-1},\\ I&=-(A^2-C^2)^{\frac{1}{2}}[\tfrac{1}{2}\sin^{-1}(C/A)+C_1]\\ &\quad+\tfrac{1}{2}C(A^2-C^2)^{-1}[1-\ln|4(A^2-C^2)^{\frac{1}{2}}|]. \end{aligned}$$

Equations (12) together with (6) are for particles on the separatrix. For values of x so large that terms of the order of $x^{2/27}$ can be neglected as compared to terms of the order of $x^{2/9}$, we find

$$u=\tfrac{1}{2}(Kx-Mx^{\frac{1}{3}})^{\frac{3}{2}} \tag{23}$$

where K and M are constants given by

$$K=6(A/2)^{\frac{1}{2}},\quad M=\tfrac{3}{4}A(18A)^{1/6}.$$

RESULTS OF CALCULATIONS

Given here are the results of calculations for the third section of the Purdue linear electron accelerator. The equations developed above were used for these calculations. The curves of Caplan and Akeley[1] for the first two sections are extended through the third section giving the energy distribution for the accelerator. The

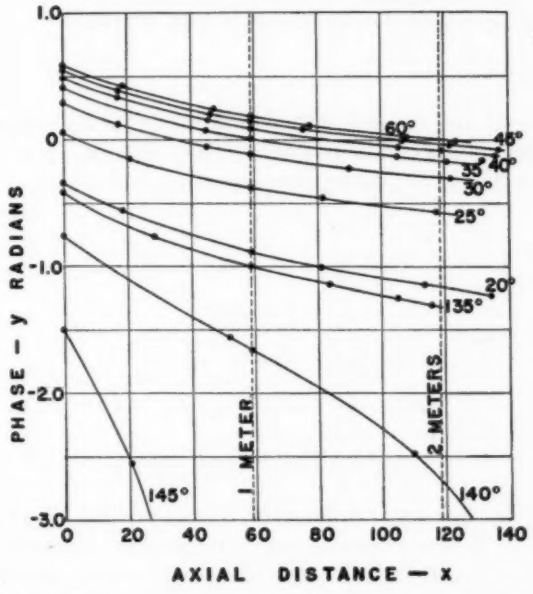


FIG. 4. Phase in radians *vs* axial distance in units of c/ω in the third section for various initial phases in the first section.

initial kinetic energy taken by them was 50 kev for each of the curves. Various initial phases were used. In Figs. 1–5 the initial phase in degrees is given by each curve.

In the solution of the differential equations, A was treated as a constant. However A will vary slowly because of the attenuation of the maximum field strength. In the calculations, the arithmetic average of

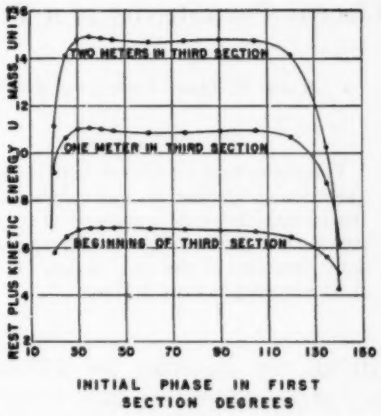


FIG. 5. Rest plus kinetic energy in mass units at various positions of the third section *vs* the initial phase in the first section. The initial energy in the first section is 50 kev.

A was taken for each meter. The wavelength of the wave was taken as 10.600 cm while the value used for E at the beginning of the section was 23.2 kev/cm with an attenuation of 10 percent per meter. From these, one obtains the values of A in the first and second meter of the section as $A_1=0.0728$ and $A_2=0.0654$. Numerical calculations were carried out for the 45° curve using an exponentially varying A instead of a constant one. The error involved in the energy, resulting from using the constant value of A, was for this case less than 0.3 percent of the total energy change for the first meter.

Since the third section will end at approximately two meters, the upper curve of Fig. 5 gives the energy of the emergent electrons as a function of the initial phase. It is seen that the maximum kinetic energy is 7.1 Mev. Figure 5 also shows that 33 percent of the electrons are in the energy range from 3.1 to 7.1 Mev, while 18 per cent are from 7.0 to 7.1 Mev. This energy distribution is for a constant phase distribution at the beginning of the first section.

The Electric Permittivity of a Dilute Suspension of Membrane-Covered Ellipsoids*

HUGO FRICKE
Walter B. James Laboratory of Biophysics, Biological Laboratory, Cold Spring Harbor, New York
(Received January 5, 1953)

This theoretical paper deals with the low frequency permittivity (ϵ) of a dilute suspension (conductivity σ, volume concentration ρ), in a conducting medium (conductivity σ_s, permittivity ϵ_s), of ellipsoidal particles (major axis $2a$) each composed of a conducting core covered with a nonconducting surface membrane (thickness t, permittivity ϵ_m). The quantity $[\epsilon-(\sigma/\sigma_s)\epsilon_s]/(\rho a\epsilon_m/t)$ is expressed in terms of Weierstrass' elliptic functions of the axial ratios, for parallel and random arrangement of the ellipsoidal axis. A table of its numerical values is given.

BIOLOGICAL cell materials are well-known examples[1,2] of an interesting dielectric mechanism by the operation of which certain types of colloidal systems may exhibit very high permittivities. The effect is obtained when conducting particles are dispersed in a conducting medium and thin insulating membranes are present at the interphases. The system acts then as a parallel and series combination of electric condensers of high capacity. The mechanism has recently attracted attention as possibly accounting for the high permittivities of certain technically important ferrites.[3,4] In connection with my early work on this effect in mammalian erythrocytes,[1,2] the theoretical relation between the permittivity of such a system and the membrane capacity was derived for particles of ellipsoidal form covered with a membrane of uniform thickness.[5] The integration in the expression obtained was carried out only for ellipsoids of rotation. In this report the calculation is carried through for the general case of ellipsoid. The general solution was needed in a study of the red blood cells of nonmammalian vertebrates which are practically perfect ellipsoids with all axis different.

The treatment is limited to suspensions in which the number of dispersed particles is so small that the local field can be taken as equal to the external field. A suspended particle is composed of a well-conducting ellipsoidal core (axis $2a\geqq 2b\geqq 2c$) covered with a membrane of thickness t_m and permittivity ϵ_m. The conductance and admittance of the membrane are supposed to be so small compared to the conductances of the particle interior and suspending medium, that practically, (1) the interior of the particle is at a constant potential and (2) the field distribution in the external medium around the particle is the same as if the whole particle were a perfect insulator.

Under these conditions, the following expression was derived for the permittivity ϵ of the suspension, for the case that the ellipsoids are arranged parallel to each other, with axis $2\alpha(\alpha=a, b \text{ or } c)$ in the direction of the external field.[5,6] Conductivity and permittivity of suspension and suspending medium are defined by writing their complex conductivities $\sigma+j(\epsilon/2)n$ and $\sigma_s+j(\epsilon_s/2)n$, respectively ($n$: cycles per second):

$$\epsilon-\epsilon_s\frac{\sigma}{\sigma_s}=\rho\frac{\epsilon_m}{t_m}ag_a, \tag{1}$$

$$g_a=\frac{4}{(2-abcL_a)^2}\cdot\frac{\int x^2\cdot dS}{\frac{4}{3}\pi a^2bc}, \tag{2}$$

$$L_a=\int_0^\infty\frac{d\lambda}{(\alpha^2+\lambda)[(a^2+\lambda)\cdot(b^2+\lambda)\cdot(c^2+\lambda)]^{\frac{1}{2}}}, \tag{3}$$

ρ is the fractional volume of the suspended ellipsoids, and x is the distance of a membrane element dS from a plane (yz) through the center of the ellipsoid, at right angle to the electric field F. The quantity $2Fx/(2-abcL_a)$ is the potential across dS and Eq. (1) expresses the fact that the electrostatic energy stored in the equivalent capacity of the suspension is equal to that stored at the membranes and in the suspending medium.

The value of ϵ for a suspension of ellipsoids of random orientation is obtained from Eq. (1) by replacing g_a by

$$g=\tfrac{1}{3}\cdot(g_a+g_b+g_c). \tag{4}$$

(1) The integral $\int x^2dS$, which represents the moment of inertia of the surface of the ellipsoid with respect to the yz plane, can be expressed as follows, in terms of Weierstrass' elliptic functions.[7] A, B, and C are the values of $\int x^2dS$, when the electric field is parallel to $2a$, $2b$, and $2c$, respectively ($j=(-1)^{\frac{1}{2}}$):

$$A=-\frac{2j}{b^2c^2(a^2-b^2)(a^2-c^2)}\cdot(a^4D_1-a^2D_2+D_3), \tag{5}$$

$$B=-\frac{2j}{a^2c^2(b^2-c^2)(b^2-a^2)}\cdot(b^4D_1-b^2D_2+D_3), \tag{6}$$

$$C=-\frac{2j}{b^2a^2(c^2-a^2)(c^2-b^2)}\cdot(c^4D_1-c^2D_2+D_3), \tag{7}$$

* Supported by the U. S. Office of Naval Research.
[1] H. Fricke, J. Gen. Physiol. **9**, 137 (1925).
[2] H. Fricke and S. Morse, J. Gen. Physiol. **9**, 153 (1925).
[3] E. Blechschmidt, Physik. Z. **39**, 212 (1938).
[4] C. G. Koops, Phys. Rev. **83**, 121 (1951).
[5] H. Fricke, Phys. Rev. **26**, 678 (1925).
[6] H. Fricke, Phys. Rev. **24**, 575 (1924).
[7] A. H. H. Tallqvist, Finska Vetenskaps—soc. Acta Soc. Sci. Fennicae. Helsingfors. **17**, 491 (1891).

$$D_1=-\frac{8\pi j}{(\wp'(\omega_0))^3}\{\zeta(\omega_0)+\omega_0\wp(\omega_0)\}, \tag{8}$$

$$D_2=\frac{4\pi j}{(\wp'(\omega_0))^5}\{3\zeta(\omega_0)\wp''(\omega_0)+\omega_0[3\wp''(\omega_0)\wp(\omega_0)-(\wp'(\omega_0))^2]\}, \tag{9}$$

$$D_3=\frac{4\pi j}{(\wp'(\omega_0))^5}\{-12\zeta(\omega_0)\wp(\omega_0)-\wp'(\omega_0)+\omega_0[\wp''(\omega_0)-12(\wp(\omega_0))^2]\}, \tag{10}$$

$$\wp(\omega_0)=-\frac{1}{3}\left(\frac{1}{a^2}+\frac{1}{b^2}+\frac{1}{c^2}\right), \tag{11}$$

$$\wp'(\omega_0)=-\frac{2j}{abc}, \tag{12}$$

$$\wp''(\omega_0)=2\left(\frac{1}{a^2b^2}+\frac{1}{b^2c^2}+\frac{1}{a^2c^2}\right), \tag{13}$$

$$\omega_0=j\frac{F(\arccos c/a,\ k)}{(a^2-c^2)^{\frac{1}{2}}}ac. \tag{14}$$

F: Elliptic integral of first kind,

$$\text{modulus } k=\frac{c(a^2-b^2)^{\frac{1}{2}}}{b(a^2-c^2)^{\frac{1}{2}}}. \tag{15}$$

The zeros of S are $e_1=\frac{1}{3}(2/c^2-1/a^2-1/b^2)$; $e_2=\frac{1}{3}(2/b^2-1/a^2-1/c^2)$; $e_3=\frac{1}{3}(2/a^2-1/b^2-1/c^2)$ by means of which $\zeta(\omega_0)$ was calculated from Weierstrass series.[8]

The expressions degenerate when two of the axis are equal. This is the case which was considered before.[5] Since there are two typographical errors in the expressions given in this reference, the correct forms are given here.

$a=b$:

$$A=B=\pi a^4\left\{1+\left(\frac{c}{a}\right)^2\frac{1}{e}\ln\left[\frac{a}{c}(1+e)\right]\right\}-\frac{\pi}{2}a^4\left\{1+\frac{1}{2}\left(\frac{c}{a}\right)^2\frac{1}{e^2}-\frac{1}{2}\left(\frac{c}{a}\right)^4\frac{1}{e^3}\ln\left[\frac{a}{c}(1+e)\right]\right\}, \tag{16}$$

$$C=\pi a^2c^2\left\{1+\left(\frac{c}{a}\right)^2\frac{1}{2e^2}-\left(\frac{c}{a}\right)^4\frac{1}{2e^3}\ln\left[\frac{a}{c}(1+e)\right]\right\}. \tag{17}$$

[8] E. Jahnke und F. Emde, *Funktionentafeln mit Formlen und Kurven* (B. G. Teubner, Leipzig, 1909).

TABLE I. Values of g's for different axial ratios.

a/b	b/c	g_a	g_b	g_c	g	a/b	b/c	g_a	g_b	g_c	g
1	1	2.25	2.25	2.25	2.25	3	1	2.51	0.98	0.98	1.49
1	1.5	2.12	2.12	1.93	2.06	3	1.5	2.90	0.84	0.97	1.57
1	2	2.13	2.13	1.85	2.04	3	2	3.33	0.76	1.04	1.71
1	3	2.29	2.29	1.98	2.19	3	3	4.45	0.77	1.21	2.14
1	4	2.55	2.55	2.22	2.44	3	4	5.65	0.88	1.40	2.64
1	6	3.17	3.17	2.76	3.03	3	6	8.08	1.14	1.79	3.67
1.5	1	2.08	1.64	1.64	1.79	4	1	2.96	0.78	0.78	1.51
1.5	1.5	2.11	1.40	1.61	1.71	4	1.5	3.52	0.62	0.79	1.64
1.5	2	2.27	1.34	1.66	1.76	4	2	4.14	0.57	0.83	1.85
1.5	3	2.75	1.48	1.86	2.03	4	3	5.54	0.59	0.96	2.36
1.5	4	3.33	1.72	2.11	2.39	4	4	7.12	0.67	1.13	2.97
1.5	6	4.52	2.19	2.65	3.12	4	6	10.4	0.85	1.46	4.24
2	1	2.14	1.33	1.33	1.60	6	1	4.01	0.54	0.54	1.70
2	1.5	2.30	1.14	1.29	1.58	6	1.5	4.81	0.43	0.56	1.93
2	2	2.58	1.06	1.39	1.68	6	2	5.82	0.39	0.59	2.27
2	3	3.33	1.13	1.62	2.03	6	3	8.12	0.41	0.68	3.07
2	4	4.10	1.31	1.85	2.42	6	4	10.5	0.45	0.80	3.92
2	6	5.63	1.68	2.32	3.21	6	6	15.4	0.57	1.04	5.67

$b=c$:

$$A=\pi a^2b^2\left(1-\frac{1}{2e^2}+\frac{a}{2be^3}\arcsin e\right), \tag{18}$$

$$B=C=\pi b^4\left(1+\frac{a}{be}\arcsin e\right)-\frac{\pi}{2}b^4\left(1-\frac{1}{2e^2}+\frac{a}{2be^3}\arcsin e\right), \tag{19}$$

$$e=\left(1-\left(\frac{c}{a}\right)^2\right)^{\frac{1}{2}}.$$

(2) By substituting

$$\lambda=\frac{a^2-c^2}{x^2}-a^2 \quad\text{and}\quad \lambda=\frac{b^2x^2-c^2}{1-x^2}$$

in L_a and L_c, respectively, these integrals are reduced to elliptic integrals of first and second kind:

$$abcL_a=\frac{2abc}{(a^2-b^2)(a^2-c^2)^{\frac{1}{2}}}\cdot\left(F\left(\arccos\frac{c}{a},\ k\right)-E\left(\arccos\frac{c}{a},\ k\right)\right), \tag{20}$$

$$abcL_c=\frac{2a^2}{a^2-c^2}-\frac{2abc}{(b^2-c^2)(a^2-c^2)^{\frac{1}{2}}}\cdot\left(E\left(\frac{\pi}{2},\ k\right)-E\left(\arcsin\frac{c}{b},\ k\right)\right). \tag{21}$$

$$\text{Modulus } k=\left(\frac{a^2-b^2}{a^2-c^2}\right)^{\frac{1}{2}}.$$

L_b can thereafter be calculated, since

$$abcL_b=2-abcL_a-abcL_c. \tag{22}$$

These expressions degenerate also when two of the axis are equal. The solution of this case was given in.[6]

Numerical values of the g's are recorded in Table I.

For a parallel array of ellipsoids of rotation placed in a field perpendicular to the axis, we obtain in the limiting case $a \to \infty$,

$$\epsilon - \epsilon_s(\sigma/\sigma_s) = \frac{9\pi}{8}\frac{\epsilon_m}{t} b\rho.$$

The numerical factor is not greatly different from that in the expression

$$\epsilon - \epsilon_s(\sigma/\sigma_s) = 4\frac{\epsilon_m}{t} b\rho$$

obtained for particles of circular cylindrical form (radius b) similarly arranged.

JOURNAL OF APPLIED PHYSICS VOLUME 24, NUMBER 5 MAY, 1953

Radiation Resistance of a Small Circular Loop in the Presence of a Conducting Ground*

JAMES R. WAIT
Radio Physics Laboratory, Defense Research Board, Ottawa, Canada
(Received January 8, 1953)

The solution is given for the total power flow from an oscillating magnetic dipole, situated above a flat homogeneous conducting ground and oriented in the vertical direction. The result is employed to derive an expression for the radiation resistance of a small wire loop whose axis is perpendicular to the ground. It is pointed out that for a finitely conducting ground that the radiation resistance is very large when the height of the small loop or dipole is a small fraction of a wavelength.

INTRODUCTION

IT is well known that the ground has a pronounced effect on the efficiency of radiating systems. Sommerfeld and Renner[1] have investigated the power flow from electric dipole type antennas over a flat conductive earth. They have shown that the radiation resistance of the dipole, as calculated for the case of an infinite insulating space, is modified considerably by the presence of the ground and the values of its electrical constants. In this paper the corresponding problem for a vertical magnetic dipole-type antenna is considered. The dipole is equivalent to a small circular loop of area dA, carrying a circulating current I that varies as $\exp(i\omega t)$, with its axis perpendicular to the ground.

A conventional cylindrical coordinate system (ρ, ϕ, z) is chosen such that the ground corresponds to negative values of z and the air corresponds to positive values of z. The conductivity and dielectric constant of the ground are denoted by σ and ϵ_2, respectively, and the dielectric constant of the air is denoted by ϵ_1. The magnetic permeability of the whole space is taken as μ and equal to that of free space. The dipole is situated at $z=h$ and is oriented in the z direction. Expressions for the total power flow from the dipole are now derived.

SOLUTION OF PROBLEM

The formal solution of this problem was given as a special case in a previous analysis by the author.[2] The magnetic field components can be written for $z \geq 0$ in terms of a scalar function F as follows:

$$i\mu\omega H_\rho = \partial^2 F/\partial\rho\partial z, \quad H_\phi = 0,$$

and

$$i\mu\omega H_z = \left(\beta^2 + \frac{\partial^2}{\partial z^2}\right)F, \tag{1}$$

where

$$\beta^2 = \epsilon_1\mu\omega^2.$$

The electric field components are given by

$$E_\rho = E_z = 0, \qquad E_\phi = \partial F/\partial\rho. \tag{2}$$

The scalar function F for this particular case for $z \geq 0$ was shown previously[3] to be

$$F = iK\int_0^\infty J_0(\lambda\rho)e^{\pm(h-z)u_1}\lambda u_1^{-1}d\lambda, \qquad -iK\int_0^\infty J_0(\lambda\rho)e^{-(z+h)u_1}G\lambda u_1^{-1}d\lambda, \tag{3}$$

where

$$G = \frac{u_2 - u_1}{u_2 + u_1}, \qquad K = \frac{\mu\omega I dA}{4\pi},$$

$$u_1 = (\lambda^2 - \beta^2)^{\frac{1}{2}}, \quad u_2 = (\lambda^2 + \gamma^2)^{\frac{1}{2}},$$

and $\gamma^2 = i\sigma\mu\omega - \epsilon_2\mu\omega^2$. The positive sign in the exponent in the first integral of the preceding equation is taken for $z > h$ and the negative sign is taken for $0 \leq z < h$. J_0 is the Bessel function of the first type of zero order.

The time-average power flow in the z direction can be calculated from the z component P_z of the Poynting

* This work was carried out under Radio Physics Laboratory Project No. 5519.
[1] A. Sommerfeld, and F. Renner, Ann. Phys. **41**, 1–36, 1942.
[2] J. R. Wait, Can. J. Phys. **29**, 577–592, 1951.

vector and is given by

$$P_z=\tfrac{1}{2}\,\mathrm{Re}\,E_\phi H_\rho^*, \tag{4}$$

where Re means the "real part of." The total power flow through a horizontal plane at $z=h+\delta$, is now denoted by S_+ and the total power flow through a horizontal plane at $z=h-\delta$ is denoted by S_- so that

$$S_\pm=2\pi\int_0^\infty P_z\rho d\rho\Big]_{z=h\pm\delta}. \tag{5}$$

Now the H_ρ and E_ϕ field components are given by

$$H_\rho=\frac{K}{\mu\omega}\int_0^\infty[\pm e^{\pm(h-z)u_1}-Ge^{-(h+z)u_1}]\lambda^2J_1(\lambda\rho)d\lambda, \tag{6}$$

and

$$E_\phi=-iK\int_0^\infty[e^{\pm(h-z)u_1}-Ge^{-(h+z)u_1}]\lambda^2J_1(\lambda\rho)u_1^{-1}d\lambda, \tag{7}$$

where again the + sign is taken for $z>h$ and the − sign for $z<h$. For convenience the distance δ is taken to be very small compared to h.

The total power flow is then given by the following triple integral

$$S_\pm=\mathrm{Re}\,\frac{iK^2\pi}{\mu\omega}\int_0^\infty\int_0^\infty\int_0^\infty f_2(\lambda)f_1^*(m)\lambda^2m^2J_1(\lambda_\rho)\times J_1(m\rho)\rho d\rho dm d\lambda, \tag{8}$$

where

$$f_1^*(\lambda)=\pm\exp(-\delta u_1^*)-G^*\exp(-2hu_1^*),$$

$$f_2(\lambda)=e^{-\delta u_1}-Ge^{-2hu_1}.$$

The variable of integration corresponding to λ in Eq. (6) is replaced by m, before introducing it into Eq. (4) to avoid confusion with the variable λ arising from Eq. (7). Now, as Niessen[3] points out, triple integrals of this type can be reduced to a single integral by making use of the orthogonal integral property of the Bessel function, namely that

$$\int_0^\infty J_1(\lambda\rho)J_1(m\rho)\rho d\rho=\begin{cases}0 & \text{if } m\neq\lambda,\\ \infty & \text{if } m=\lambda,\end{cases} \tag{9}$$

so that for infinitely small η

$$\int_{\gamma-\eta}^{\lambda+\eta}\int_{\rho=0}^\infty J_1(\lambda\rho)J_1(m\rho)\rho d\rho m dm=1, \tag{10}$$

Employing this convenient result the expression for the power flow reduces to

$$S_\pm=\mathrm{Re}\,\frac{iK^2\pi}{\mu\omega}\int_0^\infty\lambda^3f_1^*(\lambda)f_2(\lambda)d\lambda. \tag{11}$$

[3] K. F. Niessen, Ann. phys. 37, 444–454 (1938).

Now obviously the total power S supplied by the dipole to its surroundings is simply given by

$$S=S_+-S_-.$$

Without difficulty, S can now be written in terms of three integrals as follows.

$$S=\frac{2\pi K^2}{\mu\omega}[I_0-I_0'+I_c], \tag{12}$$

where

$$I_0=\mathrm{Re}\,i\int_0^\infty\lambda^3u_1^{-1}e^{\delta(u_1+u_1^*)}d\lambda, \tag{13}$$

$$I_0'=\mathrm{Re}\,i\int_0^\infty\lambda^3u_1^{-1}e^{-2hu_1}d\lambda, \tag{14}$$

and

$$I_c=\mathrm{Re}\,2i\int_0^\infty\lambda^3(u_1+u_2)^{-1}e^{-2hu_1}d\lambda, \tag{15}$$

and where it has been assumed that $\delta\ll h$.

Each of these integrals can be clearly interpreted. The first one I_0 corresponds to the power flow (apart from a constant factor) into an infinite insulating space. The second one I_0' accounts for a geometrical image of the dipole in the ground plane. The third one I_c can be interpreted as a correction to account for the finite conductivity of the ground and it approaches zero as the propagation constant γ approaches infinity. Of course, both the second and third integrals become negligible as h approaches large values.

These integrals are now evaluated. In the integral I_0 it is observed that the integrand is real for $\lambda>\beta$ and that $u_1+u_1^*=0$, so

$$I_0=\mathrm{Re}\,i\int_0^\beta\lambda^3(\lambda^2-\beta^2)^{-\frac{1}{2}}d\lambda=2\beta^3/3. \tag{16}$$

Clearly the second integral can be written

$$I_0'=\mathrm{Re}\,i\int_0^\infty(u_1^2+\beta^2)e^{-xu_1}\lambda u_1^{-1}d\lambda, \tag{17}$$

where $x=2h$. This can now be expressed as

$$I_0'=\mathrm{Re}\,i\left(\frac{\partial^2}{\partial x^2}+\beta^2\right)\int_0^\infty e^{-xu_1}\lambda u_1^{-1}d\lambda, \tag{18}$$

which is easily evaluated to give

$$I_0'=\mathrm{Re}\,I\left(\frac{\partial^2}{\partial x^2}+\beta^2\right)\frac{e^{-i\beta x}}{x}=\frac{2}{x^3}(\sin\beta x-\beta x\cos\beta x), \tag{19}$$

The third integral is more difficult to evaluate; it can, however, be written

$$I_c=\mathrm{Re}\,i2\left(\frac{\partial^2}{\partial x^2}+\beta^2\right)J_c, \tag{20}$$

where

$$J_e=\int_0^\infty e^{-u_1 z}(u_1+u_2)^{-1}\lambda d\lambda. \quad (21)$$

This latter is then broken into three separate integrals as follows

$$(\gamma^2+\beta^2)J_e=\int_0^\infty \gamma e^{-u_1 z}\lambda d\lambda-\int_0^\infty u_1 e^{-u_1 z}\lambda d\lambda + \int_0^\infty (u_2-\gamma)e^{-u_1 z}\lambda d\lambda. \quad (22)$$

The first two integrals can be readily evaluated in a similar manner to I_0'. The third can be evaluated only approximately under the conditions that $|\gamma x|$ is much greater than unity, such that it is permissible to replace

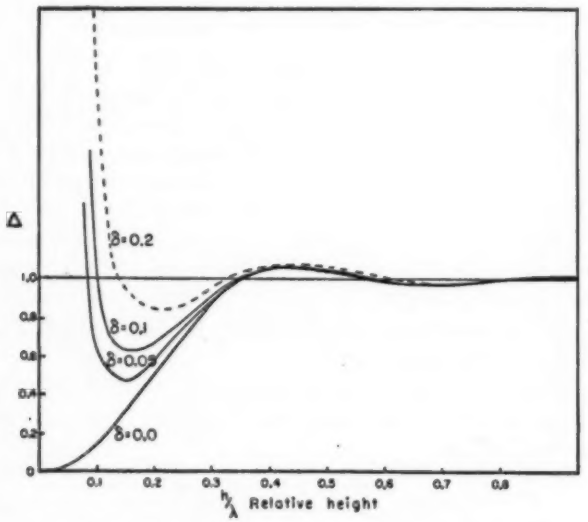


FIG. 1. The ratio of the radiation resistance of the loop over a conducting ground to the radiation resistance in free space.

$u_2-\gamma$ by $\lambda^2/2\gamma$ in the integrand. Carrying out these steps it is found that

$$x^3(\gamma^2+\beta^2)e^{+i\beta z}J_e=(\gamma x)^{-1}(1+i\beta x) -(\gamma x)^{-2}(2+2i\beta x-\beta^2x^2)+(\gamma x)^{-3}(3+3i\beta x-\beta^2x^2). \quad (23)$$

The three terms in the right-hand side of the preceding equation correspond to the values of the three integrals in Eq. (22). The approximation made in the latter integral means that terms of order $(\gamma x)^{-4}$ and higher have been neglected.

The expression for I_e is then obtained by carrying out the operations indicated in Eq. (20). Doing this and then retaining only terms of order $(\gamma x)^{-1}$, there results

$$I_e\approx \mathrm{Re}\, 4i\gamma x^{-4}(\gamma^2+\beta^2)^{-1}(3+3i\beta x-\beta^2x^2)e^{-i\beta z}. \quad (24)$$

Now the displacement currents in the ground are small compared to the conduction currents[2] at low radio frequencies since $\sigma\gg\epsilon_2\omega$ so that $\gamma\approx m/(1-i)$, where $m=(2\sigma\mu\omega)^{\frac{1}{2}}$. If this approximation is valid then also $\beta^2\ll|\gamma^2|$, so that finally

$$I_e\approx 4m^{-1}x^{-4}[(3-3\alpha-\alpha^2)\cos\alpha + (3+3\alpha-\alpha^2)\sin\alpha]. \quad (25)$$

The total effective radiation resistance R of the magnetic dipole in the presence of the ground is then defined by

$$R=2S/|I^2|, \quad (26)$$

Combining the preceding results it follows that

$$R=\frac{\mu\omega\pi a^4}{4}(r_0-r_0'+r_e), \quad (27)$$

where

$$r_0=2\alpha^3/3x^3,$$
$$r_0'=2x^{-3}(\sin\alpha-\alpha\cos\alpha),$$
$$r_e=4m^{-1}x^{-4}[(3-3\alpha-\alpha^2)\cos\alpha+(3+3\alpha-\alpha^2)\sin\alpha],$$

and where $\alpha=\beta x=2\beta h$.

The significance of this result is best seen by defining a factor Δ by

$$\Delta=(r_0-r_0'+r_e)/r_0, \quad (28)$$

which actually is the ratio of the value for the finitely conducting ground to the corresponding value in free space. The factor Δ is plotted as a function h/λ in Fig. 1, where λ is the wavelength in air given by

$$\lambda=2\pi/\beta.$$

Curves are plotted for three values of a loss factor δ which is defined by

$$\delta=\beta/m=(\epsilon_1\omega/2\sigma_2)^{\frac{1}{2}}.$$

From these curves it is seen that the effect of the ground is to modify appreciably the radiation resistance for heights less than about 0.6 of a free space wavelength As would be expected the radiation resistance is zero if the height of the dipole is zero and the ground is perfectly conducting. For a finitely conducting ground however the resistance becomes very large for low heights. In fact as h approaches zero the ratio Δ approaches infinity as

$$\Delta\approx 18\delta/\alpha^4.$$

That is, at very low heights the power absorbed in the ground is many times the total power radiated from the same dipole if it were at heights greater than a wavelength.

SOLUTION FOR LOW HEIGHTS

It is possible to obtain a more accurate solution for low heights by assuming at the outset that displacement currents in the air can be neglected. This is justified, as pointed out previously,[2] when all significant dimensions are much less than a free-space wavelength which is

effected by setting the propagation constant β in the air equal to zero.

With this approximation the radiation resistance R as defined earlier can be written

$$R=\text{Re } i\mu\omega\pi a^4 Q/4, \tag{29}$$

where

$$Q=\int_0^\infty \lambda^2\left[\frac{\lambda-(\lambda^2+\gamma^2)^{\frac{1}{2}}}{\lambda+(\lambda^2+\gamma^2)^{\frac{1}{2}}}\right]e^{-\lambda b}d\lambda, \tag{30}$$

with $b=2h$. This can be rewritten in the following manner:

$$Q=\frac{1}{\gamma^2}\left[\int_0^\infty 2\lambda^3(\lambda^2+\gamma^2)^{\frac{1}{2}}e^{-\lambda b}d\lambda - \int_0^\infty \lambda^2(2\lambda^2+\gamma^2)e^{-\lambda b}d\lambda\right]. \tag{31}$$

The integrals are reduced to a simpler form by expressing them as derivatives with respect to b of two new integrals as follows:

$$Q=-\frac{2}{\gamma^2}\left[\frac{\partial^3}{\partial b^3}\int_0^\infty (\lambda^2+\gamma^2)^{\frac{1}{2}}e^{-\lambda b}d\lambda + \left(\frac{\gamma^2}{2}\frac{\partial^2}{\partial b^2}+\frac{\partial^4}{\partial b^4}\right)\int_0^\infty e^{-\lambda b}d\lambda\right]. \tag{32}$$

The following integral given by Watson[4] can now be employed:

$$\int_0^\infty (1+g^2)^{\frac{1}{2}}e^{-\alpha g}dg=\frac{\pi}{2\alpha}[H_1(\alpha)-Y_1(\alpha)], \tag{33}$$

where $H_1(\alpha)$ and $Y_1(\alpha)$ are Struve's function and Bessel's function of the second type, respectively, for an argument α. Employing this integral, the expression for Q is given by

$$Q=-\frac{2}{b^3}\left[1+\frac{24}{\alpha^2}+\alpha^3\frac{\partial^3 T(\alpha)}{\partial\alpha^3}\right], \tag{34}$$

[4] G. N. Watson, *Theory of Bessel Functions* (Cambridge University Press, Cambridge, 1952), second edition, p. 332.

where

$$T(\alpha)=\frac{\pi}{2\alpha}[H_1(\alpha)-Y_1(\alpha)],$$

and $\alpha=\gamma b$. Asymptotic expansions for the functions $T(\alpha)$ have been developed[4] and by employing this result it is easily shown that

$$Q=-\frac{2}{b^3}\left[1+\frac{24}{\alpha^2} \sum_{m=0}^{p-1}\frac{(2m+1)(2m+2)(2m+3)\Gamma(m+\frac{1}{2})2^{2m-1}}{\Gamma(\frac{3}{2}-m)\alpha^{2m+1}}+R_p\right], \tag{35}$$

where the magnitude of the remainder R_p, is of the order of α^{-2p}.

The expression for the radiation resistance is then

$$R=\frac{3\mu\omega\pi a^4}{16h^3}\left[\frac{1}{mh}-\frac{4}{(mh)^2}+\frac{5}{(mh)^2}-R_4\right], \tag{36}$$

where the remainder R_4 is of the order of $(mh)^{-4}$. The value of R to the order of $(mh)^{-1}$ agrees with the result given by Eq. (27) for the case when βh is much less than unity. In Eq. (36) m is defined by $m=(2\sigma_2\mu\omega)^{\frac{1}{2}}$.

CONCLUSIONS

It is quite apparent that the radiation resistance of a magnetic dipole situated over a conducting ground, has a dependence on the electrical constants of the ground. For heights less than a wavelength this effect becomes very pronounced. At very small heights the radiation resistance becomes very large. No doubt, the possibility exists of utilizing this property in devising a method to measure the value of the electrical conductivity of the ground at low radio frequencies from a loop or coil fixed on an aircraft.

ACKNOWLEDGMENT

Dr. W. Petrie read the manuscript and made several suggestions.

A.S.E.E. Convention—University of Florida June 22–26, 1953

The 61st annual convention of the American Society for Engineering Education will be held next June 22–26 inclusive at the University of Florida in Gainesville. This year's host institutions—the University and the State of Florida—are making every effort to arrange a technical program of unusual merit, coupled with a recreational program of the type that has made Florida a favorite vacationland.

In addition to an outstanding technical program in the field of engineering education, considerable attention will be given to the currently vital problems of manpower utilization and industrial expansion.

Dynamic Mechanical Properties of Polyisobutylene

EDWIN R. FITZGERALD,* LESTER D. GRANDINE, JR.,† AND JOHN D. FERRY
Department of Chemistry, University of Wisconsin, Madison, Wisconsin

(Received January 16, 1953)

Values of complex shear compliance (and rigidity) have been obtained for a sample of polyisobutylene of viscosity-average molecular weight 1.35×10^6 at 22 temperatures from −45 to 100°C and at about 20 frequencies from 30 to 5000 cps. Measurements were made by means of the transducer method of Fitzgerald and Ferry with a precision of ±2 percent and are estimated to be accurate to within ±3 percent as evidenced by agreement obtained between 7 different samples of widely varying dimensions. Values of the real part of the complex shear compliance J' vary from 3.1×10^{-7} cm²/dyne at 99.9°C to 1.0×10^{-10} cm²/dyne at −44.6°C. The frequency dependence of the loss tangent J''/J' indicates the presence of a low, broad maximum of 1.7 at −10 to −5°C and a second, smaller maximum at lower temperatures. This second maximum is also evident in a plot of J''/J' at a fixed frequency against temperature. The wide temperature and frequency ranges of the measurements have provided an essentially complete experimental description of the dispersion region in polyisobutylene corresponding to the transition from a rubber-like elasticity to that of a hard glass.

INTRODUCTION

SOME data on the dynamic mechanical properties of polyisobutylene of viscosity-average molecular weight 1.2×10^6 have been previously reported;[1,2] these data extended from 25 to 250 cps and over a temperature range from −15 to 50°C. The development of a new method for determining dynamic mechanical properties as described by Fitzgerald and Ferry[3] has subsequently made possible large extensions of the frequency and temperature ranges of these measurements and, at the same time, has resulted in greater precision in the determination of the complex shear compliance and rigidity. An unfractionated polyisobutylene (PIB) of viscosity average molecular weight 1.35×10^6 distributed by the National Bureau of Standards[4] was chosen for the present study. Seven samples of widely varying dimensions were used to determine the complex shear compliance (and rigidity) at 22 temperatures from −45 to 100°C and at about 20 frequencies from 30 to 5000 cps.

TABLE I. Dimensions and coefficients of samples used in measurements on polyisobutylene.

Sample No.	Mass grams m_1	Mass grams m_2	Approx. molded diam. of each disk in.	Approx. molded thickness of each disk in.	Sample coefficient C at 25°C cm	Percent compression	Bulging[a] correction percent
8	0.1722	0.1680	1 1/16	1/32	58.5	8	None
17	0.3412	0.3476	1 1/16	1/16	31.5	7	None
18	0.1758	0.1662	1 1/16	1/32	59.9	6	None
19	0.1414	0.1342	7/16	1/8	3.63	10	13
19a[b]	0.1414	0.1342	7/16	1/8	3.73	9	13
25	0.3008	0.2980	1 1/16	1/16	2.62	5	3
26	0.3018	0.2986	1 1/16	1/16	2.73	6	4

[a] For thin samples or where no bulging occurs, $C=A_1/h_1+A_2/h_2$, where A_1, A_2 are the cross-sectional areas and h_1, h_2 the thicknesses of the sample disks after compression.
[b] Sample 19 after having been removed and then reinserted in the apparatus during the course of the measurements.

MATERIAL STUDIED

The polyisobutylene measured was furnished by Dr. Robert S. Marvin of the National Bureau of Standards. Each sample consisted of two disks of nearly identical dimensions which were formed by molding for 6 to 9 hours at 85°C in a stainless steel mold. A summary of the dimensions and sample coefficients of the seven samples used is given in Table I. Values of the sample coefficients at temperatures other than 25°C were calculated from the thermal expansion coefficient of PIB, assuming that the values of thicknesses, fixed by the dimensions of the apparatus, were independent of temperature. A complete description of the method for determining sample coefficients and the effect of bulging that may occur as a result of compression has been given in a previous publication.[3]

DETERMINATION OF COMPLEX SHEAR COMPLIANCE

Measurements of complex shear compliance ($J^*=J'-iJ''$) were made by means of the transducer method of Fitzgerald and Ferry.[3] A rigid tube with coils around the ends is suspended by wires so that the coils are in radial magnetic fields. Alternating currents are passed through the coils, and the consequent motion is modified by introducing the sample in the form of two disks between the tube and a suspended inner core of large mass. From determinations of the mechanical impedance of the system with and without a sample the impedance of the sample alone is found. One transducer coil is placed in the arm of a bridge where it acts as an electrical impedance $\mathbf{Z}_2'$, while the other parallels the bridge circuit. $\mathbf{Z}_2'$ is found for various current ratios in the coils, and the mechanical impedance is obtained from the differences in electrical impedance values. While the values of electrical impedance can be

* Present address: Department of Physics, Pennsylvania State College.
† Union Carbide and Carbon Fellow in Physical Chemistry, 1950–1952. Present address: E. I. du Pont de Nemours and Company, Wilmington, Delaware.
[1] Marvin, Fitzgerald, and Ferry, J. Appl. Phys. **21**, 197 (1950).
[2] Ferry, Fitzgerald, Johnson, and Grandine, J. Appl. Phys. **22**, 717 (1951).
[3] E. R. Fitzgerald and J. D. Ferry, J. Colloid Sci. **8**, 1 (1953).
[4] R. S. Marvin, Interim Report on the Cooperative Program on Dynamic Testing, National Bureau of Standards, 1951.

determined within 0.5 percent and values for the total mechanical impedance to within less than 1.0 percent, this does not mean that the sample admittance and hence J' and J'' (or G' and G'') are determined with the same precision. In order to determine that part of the total measured mechanical impedance (or admittance) which is due to the sample, two subtractions must be made (see Fig. 1): the mechanical impedance resulting from the mass of the driving tube and its support wires $\mathbf{Z}_{M^0}$ is subtracted from the total measured impedance $\mathbf{Z}_M$ to give a net impedance $\mathbf{Z}_{AB}$; then from the corresponding net admittance $\mathbf{Y}_{AB}$ there must be subtracted a background admittance $\mathbf{Y}_{M''}$ because of the slight motion of the heavy suspended mass against which the sample is sheared. The mechanical admittance of the sample is given by $\mathbf{Y}_{M'}=\mathbf{Y}_{AB}-\mathbf{Y}_{M''}$, where $\mathbf{Y}_{M''}$ is the admittance of the floating mass M; $\mathbf{Y}_{AB}=1/\mathbf{Z}_{AB}$, and $\mathbf{Z}_{AB}=\mathbf{Z}_M-\mathbf{Z}_{M^0}$. The complex compliance $\mathbf{J}^*$ is then found from

$$\mathbf{J}^*=-i\mathbf{Y}_{M'}C/\omega, \tag{1}$$

where C is a coefficient depending on the shape and dimensions of the sample and ω is the circular frequency. The complex shear modulus is determined from $\mathbf{G}^*=1/\mathbf{J}^*$.

As a result of the subtractions mentioned there are two situations in which the results become uncertain. With a sample of low mechanical impedance (for example, PIB at high temperatures) the total impedance measured is chiefly due to the mass m of the driving tube; at high frequencies the impedance of the driving tube itself ($\sim\omega m$) predominates and completely masks the influence of the sample. On the other hand, a sample of high impedance (for example, PIB at low temperatures) will result in strong coupling between the driving tube and suspended mass so that the two will oscillate as a single unit whose motion is independent of the sample properties. This coupling will decrease at high frequencies as the impedance ($\sim\omega M$) of the suspended mass increases. Consequently, to maintain maximum precision over a range of frequencies and temperatures a number of sample sizes are used to provide optimum sample impedances (thin samples of large diameter at high temperatures and thick samples of small diameter at low temperatures). Wherever the subtractions give differences so small as to produce an uncertainty greater than ± 2 percent, measurements on a particular sample are discontinued.

A detailed description of the experimental method, temperature control system, and measuring apparatus has been given elsewhere.[3]

EXPERIMENTAL RESULTS

Measurements from 9.8 to 99.9°C were made on sample 8. Four separate runs by two observers were made at 25.0°C; the results agreed to within 2 percent. After completing measurements at 99.9°C a check run was made at 25.0°C. The values of J''/J' (dimensionless) were unchanged, and values of J'' and J' were found to be the same as those initially determined (within less than ± 2 percent). Measurements were made on sample 17 at 25.0 and 19.9°C and agreed with those on sample 8. To establish firmly the results at 25.0°C a third sample, No. 18, with the same approximate dimensions as No. 8, was measured at 25.0°C, and again the results agreed with those on Nos. 8 and 17. A thick sample, No. 19, was measured at temperatures from 25.0 to −28.5°C. Unfortunately this sample was compressed greatly in the apparatus, and considerable bulging occurred. From a determination of the radius of curvature of the bulge (assuming it to be circular) a corrected value of sample coefficient was determined,[3] and the values of J' and J'' were then found to be about 3 percent higher than corresponding values determined at 9.8, 14.8, 19.8, and 25.0°C for samples 8, 17, and 18 (values of J''/J' were the same). Since the determination of the bulge radius is inexact, the data presented for sample 19 were actually empirically corrected to agree with that of the other samples; the same correction, of course, sufficing for all temperatures. Samples 25 and 26 were thicker and of smaller diameter than the others and were used to obtain data at 4.9, −9.9, and −19.9 to −44.6°C. In order to avoid the uncertainty resulting from bulging owing to high compression, these samples were compressed only 5 to 6 percent, which proved to be sufficient to prevent slipping. Some slight degree of bulging occurred, however, and an empirical correction of 3 to 4 percent was made in the sample coefficients to obtain agreement with samples 19 and 19a at 4.9, −9.9, and −19.9°C. (While the basis for these empirical adjustments of sample coefficients is well founded, it should be pointed out that even without them the results on samples of different dimensions agree to within 3 percent). The measurements on sample 26 checked those on No. 25 within 2 percent. After measuring No. 26 at 4.9, −9.9, −19.9, −25.0, −29.8, and −34.7°C, the temperature was raised to −19.9°C and measurements at that temperature taken again; these agreed with the initial values. Measurements were then made at −44.7°C, checking those on sample 25, and then repeated at

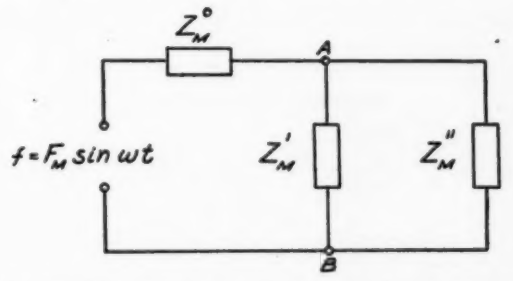


FIG. 1. Mechanical circuit diagram for transducer with sample. $\mathbf{Z}_{M^0}$ is the mechanical impedance of driving tube and support wires; $\mathbf{Z}_{M'}$ is the mechanical impedance of the sample disks; $\mathbf{Z}_{M''}$ is the mechanical impedance of the suspended heavy mass and its support wires.

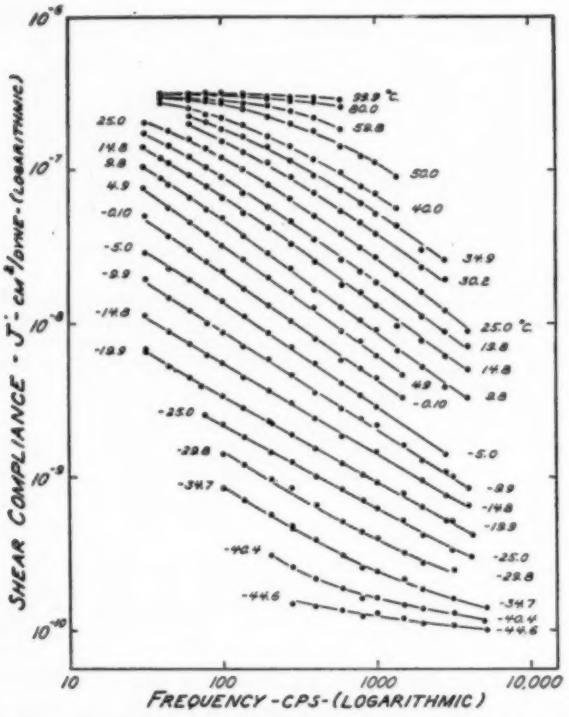


FIG. 2. Variation of the real part of the dynamic shear compliance J' with frequency for polyisobutylene ($\bar{M}_\eta = 1.35\times10^6$) at 22 temperatures as indicated.

−19.9°C, whereupon it was found that the values of J' and J'' were again the same as those initially determined at that temperature.

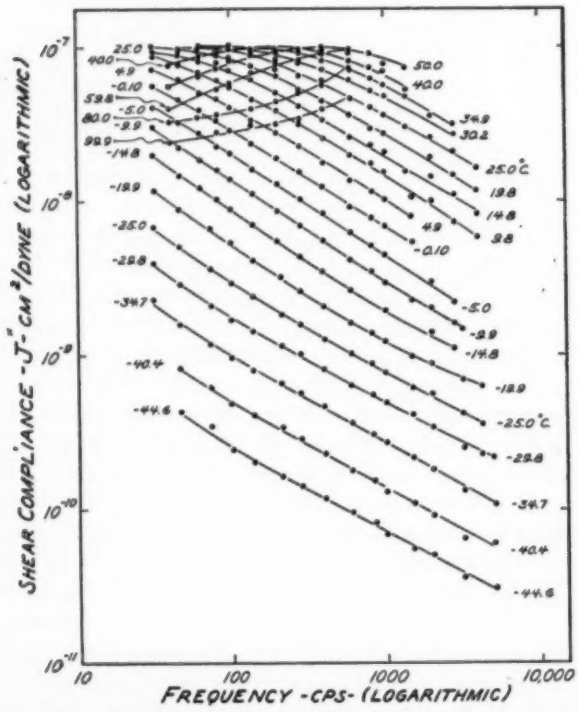


FIG. 3. Variation of the imaginary part of the dynamic shear compliance J'' with frequency for the polyisobutylene of Fig. 2.

The variation of the real part of the complex shear compliance J' with frequency at 22 temperatures is shown in Fig. 2. The values of J' vary from about 3.1×10^{-7} cm²/dyne at 99.9°C to 1.0×10^{-10} cm²/dyne at −44.6°C. At both of these temperatures the variation of J' with frequency is not very marked; the sharpest dispersion occurs in the region 0 to −15°C. The variation of the imaginary part, J'', of the shear compliance with frequency at 22 temperatures is given in Fig. 3. Here the values range from a maximum of about 1.0×10^{-7} cm²/dyne obtained at 25.0, 30.2, 34.9, 40.0, 50.0, and 59.8°C to a low value of 3.0×10^{-11} cm²/dyne at −44.6°C. The variation of the real part of the rigidity G' with frequency is shown in Fig. 4,

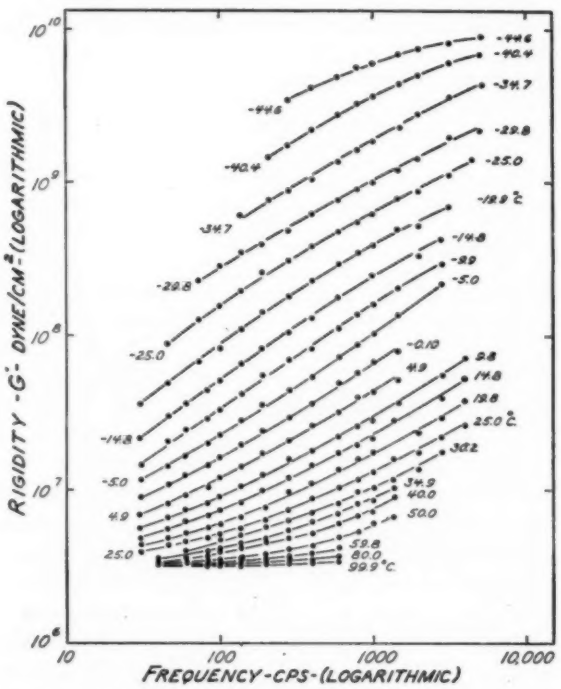


FIG. 4. Variation of the real part of the dynamic rigidity G' with frequency for the polyisobutylene of Fig. 2.

while the real part of the shear viscosity η' as a function of frequency is adduced in Fig. 5.

The variation of mechanical loss tangent J''/J' with frequency is shown in Fig. 6 for some of the temperatures at which measurements were made. The loss tangent passes through a broad maximum of about 1.7 between −15 and 0°C, but of particular interest is the second maximum, or definite plateau, that occurs in the curves at low temperatures. Since the same effect was observed for three samples of different dimensions (Nos. 19a, 25, and 26), this must be taken as a real dispersion region and not a result of some peculiarity of the experimental apparatus. This conclusion is also strengthened by the fact that the loss tangent-frequency curves for two plasticized polyvinyl chloride composi-

tions[3,5] and for polystyrene[6] showed only a single maximum when determined with this same apparatus. In each case measurements extended into regions of very low sample admittance where stray instrument effects are most likely to be present; only in the case of polyisobutylene did a second maximum appear.

The temperature variation of the real part of the shear compliance J' and the loss tangent J''/J' at a constant frequency can be determined as in Fig. 7, where values at three frequencies are plotted against temperature. Here again a second maximum in J''/J' appears at low temperatures.‡

A summary of values of J' and J'' at 22 temperatures is given in Table II in the Appendix.

DISCUSSION

There are no results previously reported on polyisobutylene with which all of the curves in Figs. 2–5 can be compared. However, the results at 25.0°C have been compared with some obtained by Hopkins and by

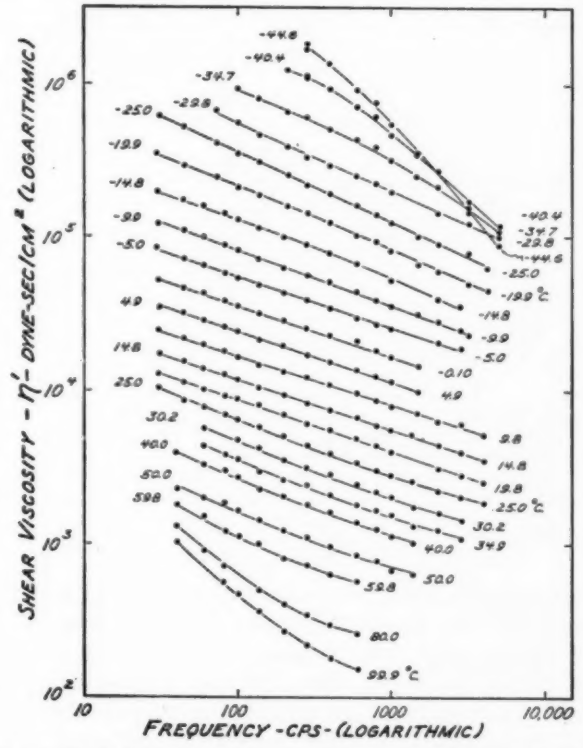


FIG. 5. Variation of the real part of the dynamic shear viscosity η' with frequency for the polyisobutylene of Fig. 2.

[5] J. D. Ferry and E. R. Fitzgerald, J. Colloid Sci. (to be published).

[6] L. D. Grandine, Jr., and J. D. Ferry, J. Appl. Phys. (to be published).

‡ *Note added in proof:* Dr. T. W. DeWitt recently reported a similar second maximum in the mechanical loss tangent *vs* temperature curves for a polyisobutylene of molecular weight two million, measured at 20, 50, and 100 cps from −60 to +20°C. His results, obtained at the Mellon Institute, are in general agreement with the date given here. (T. W. De Witt, discussion following paper GB2 at the American Physical Society Meeting, North Carolina, March 26, 1953.)

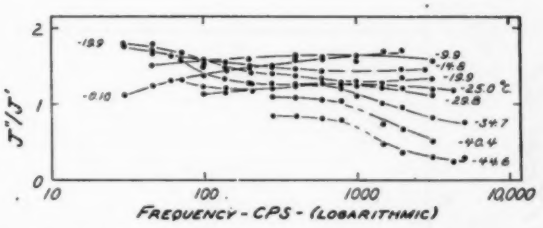


FIG. 6. Variation of mechanical loss tangent J''/J' with frequency for the polyisobutylene of Fig. 2. To avoid confusion results are given for only 9 of the 22 temperatures at which measurements were made.

Marvin and Weiss for this same sample of PIB and found to be in good agreement.[3] The general shape of the rigidity and viscosity curves as functions of frequency and the effect of temperature are also comparable with earlier results[2] obtained for a different sample of polyisobutylene of viscosity-average molecular weight 1.2×10^6.

The dispersion region for polyisobutylene as described by the curves of Figs. 2 and 3 is much broader and the variation of compliance with frequency more gradual than has been previously found for plasticized polyvinyl chloride compositions or for polystyrene. The appearance of a second maximum in the loss tangent-frequency curve (and the loss tangent-temperature curve) as shown in Figs. 6 and 7 is also in marked contrast to results for other linear polymers.[3,5,6]

An attempt to describe the effect of temperature on mechanical properties is sometimes made by plotting values of shear compliance (or rigidity) at a constant frequency against temperature as is done in Fig. 7. While such a plot does indicate the effect of temperature, it also includes a frequency dependence. This can best be illustrated by considering the effect of temperature on the distribution of retardation or relaxation times in the polymer. If all relaxation times depend identically on temperature, then it has been demonstrated[2,7] that a temperature change is practically equivalent to a frequency shift; experimental results on a number of linear polymers including polyisobutylene[8] have confirmed this view. The observed change in J', for example, at 1000 cps in going from 0 to 20°C will be a result of the temperature dependence (i.e., the change in magnitude of relaxation times from 0 to 20°C) but also the degree to which the corresponding equivalent frequency change happens to affect J'. Where the effect of frequency on J' is slight (for example, at −45 or 100°C) the temperature effect will be predominant; at other points the change in J' may result primarily from a very pronounced dependence of J' on a shift of relaxation times rather than from any great change in the relaxation times themselves. A much more instructive separation of the temperature and frequency dependence of mechanical properties can be obtained by

[7] J. D. Ferry, J. Am. Chem. Soc. 72, 3746 (1950).

[8] Ferry, Fitzgerald, Grandine, and Williams, Ind. Eng. Chem. 44, 703 (1952).

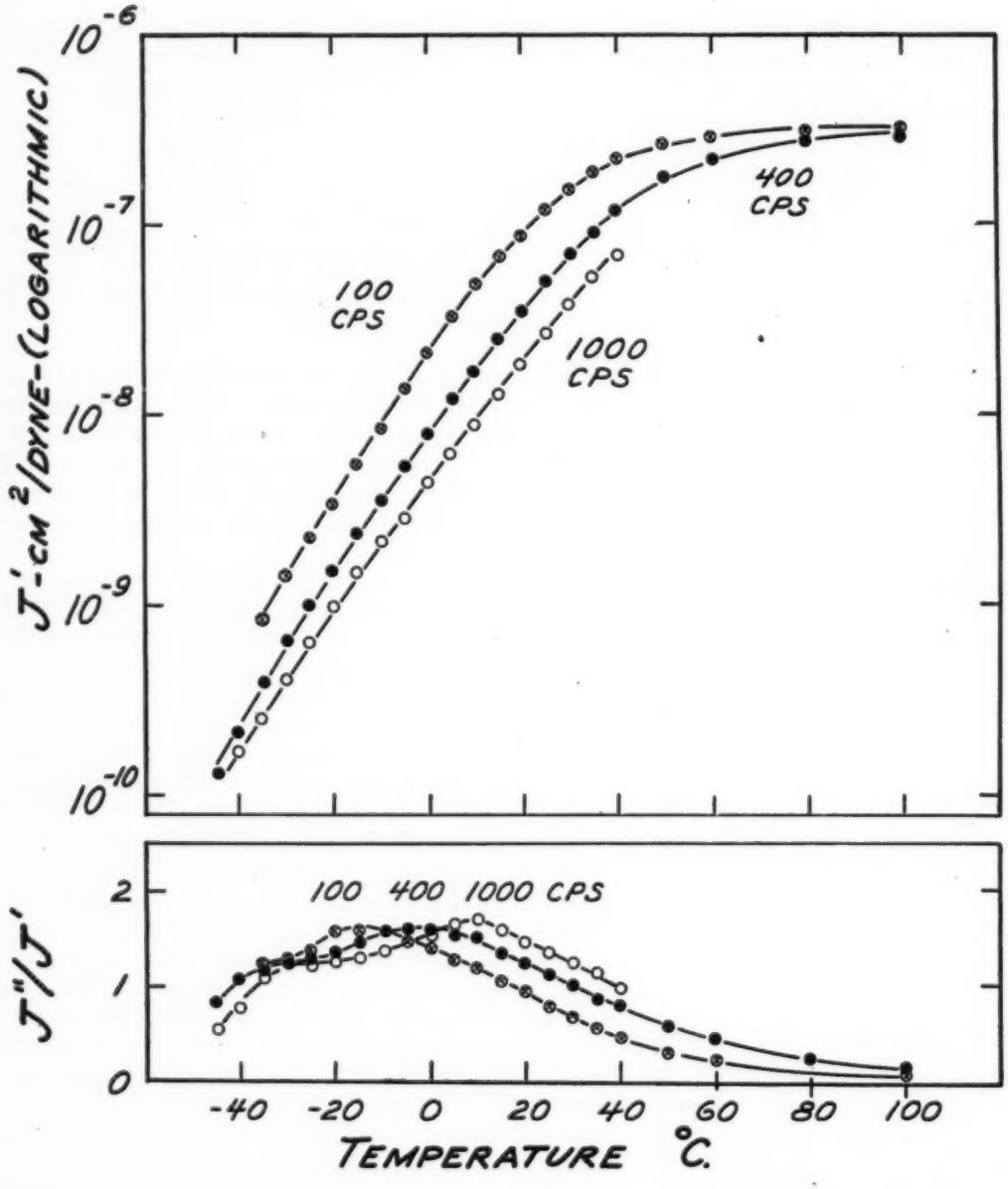


FIG. 7. Variation of the real part of the dynamic shear modulus J' and the mechanical loss tangent J''/J', with temperature for the polyisobutylene of Fig. 2, at 100, 400, and 1000 cps as indicated.

superposing data at different temperatures. This treatment as well as a more detailed discussion of the experimental results presented here will be given in a subsequent paper.

ACKNOWLEDGMENTS

We are indebted to M. L. Williams for help in carrying out some of the experimental work and for making some check measurements at low temperatures and to Miss Martha Maybury for assistance in calculations.

This work was supported in part by a grant from Research Corporation and in part by the Research Committee of the Graduate School of the University of Wisconsin from funds supplied by Wisconsin Alumni Research Foundation. Since March, 1951 it has been part of a program of research on the physical structure and properties of cellulose derivatives and other polymers supported by the Allegany Ballistics Laboratory, Cumberland, Maryland, an establishment owned by the United States Navy and operated by the Hercules Powder Company under Contract NOrd 10431. We are grateful to Union Carbide and Carbon Corporation for assistance through its Fellowship in Physical Chemistry.

APPENDIX

TABLE II. Summary of dynamic mechanical data on polyisobutylene ($\bar{M}_v = 1.35 \times 10^6$).

T °C	Frequency[a] cps	30	40	45	60	72	80	100	140	200	210	280	400	600	800	1000	1400	1500	2000	2800	3000	4000	4200	5100
−44.6	J'	...		...		...		...	...		...	0.151	0.145	0.136	0.123	0.132		0.119	0.111		0.108			0.0995
	J'' cm²/dyne(×10⁻⁹)	...		...		0.345		0.244	0.200		0.166	0.142	0.118	0.0955	0.0831	0.0697		0.0551	0.0506		0.0368			0.0309
−40.4	J'	...		...		...		...	...		0.308	0.264	0.212	0.186	0.163	0.167		0.149	0.137		0.129			0.112
	J'' cm²/dyne(×10⁻⁹)	...		0.823		0.617		0.490	0.411		0.336	0.284	0.230	0.177	0.151	0.129		0.107	0.0921		0.0651			0.0601
−34.7	J'	...		...		...		0.848	0.702		0.571	0.478	0.387	0.308	0.254	0.248		0.216	0.188		0.165			0.139
	J'' cm²/dyne(×10⁻⁹)	2.20		1.59		1.17		0.958	0.805		0.660	0.566	0.467	0.367	0.304	0.268		0.214	0.182		0.130			0.107
−29.8	J'	...		...		...		1.41	1.19		0.997	0.852	0.656	0.512	0.440	0.398		0.368	0.278		0.220		...	0.153
	J'' cm²/dyne(×10⁻⁹)	3.96		2.92		2.07		1.70	1.43		1.13	1.02	0.796	0.639	0.553	0.486		0.415	0.336		0.248		0.226	0.219
−25.0	J'	...		...		2.54		2.17	1.81		1.48	1.25	1.01	0.825	0.703	0.625		0.526	0.422		0.329		0.302	...
	J'' cm²/dyne(×10⁻⁹)	6.79		4.98		3.66		2.97	2.39		1.80	1.58	1.26	1.02	0.880	0.777		0.622	0.558		0.421		0.350	...
−19.9	J'	6.59		5.23		3.93		3.42	2.83		2.25	1.88	1.52	1.24	1.03	0.947		0.792	0.646		0.526		0.388	...
	J'' cm²/dyne(×10⁻⁹)	11.8		8.91		6.56		5.38	4.20		3.14	2.63	2.08	1.63	1.37	1.23		0.977	0.893		0.687		0.634	...
−14.8	J'	11.5		8.96	7.48		6.17	5.55	4.51	3.64		2.99	2.38	1.85	...	1.47		...	0.949	0.752		...		
	J'' cm²/dyne(×10⁻⁹)	20.1		14.8	12.2		10.2	8.76	6.93	5.47		4.42	3.49	2.65	...	1.91		...	1.40	1.10		...		
− 9.9	J'	19.7		14.8	12.0		10.0	8.70	7.06	5.87		4.71	3.55	2.78	2.43	2.18		1.62	1.25	1.08	1.05	...		
	J'' cm²/dyne(×10⁻⁹)	31.1		22.5	18.8		15.8	13.7	11.0	8.54		6.81	5.55	4.19	3.44	2.94		2.29	2.19	1.58	1.64	...		
− 5.0	J'	29.8		23.0	19.2		16.2	13.7	11.0	8.71		6.89	5.35	4.09	3.40	2.90		...	1.61	1.42		...		
	J'' cm²/dyne(×10⁻⁹)	40.4		32.7	27.9		23.2	20.4	16.3	12.9		10.4	8.39	6.21	5.15	4.43		...	3.02	2.08		...		
− 0.10	J'	50.4		37.2	30.1		25.0	21.2	16.8	13.1		10.4	7.88	5.77	4.98	4.37		3.24	...	...		...		
	J'' cm²/dyne(×10⁻⁹)	55.9		45.5	39.0		33.6	29.2	24.0	19.2		15.6	12.4	9.04	7.72	6.75		5.53	...	...		...		
4.9	J'	75.7		56.3	46.2		37.8	32.1	26.0	20.1		15.9	12.2	8.87	7.52	6.12		4.67	...	...		...		
	J'' cm²/dyne(×10⁻⁹)	72.8		61.2	53.5		46.4	40.9	33.9	27.5		23.0	17.9	13.6	11.4	10.3		8.13	...	...		...		
9.8	J'	106.		80.2	66.2		54.6	47.2	37.5	29.0		23.1	16.8	12.7	10.4	9.02		6.68	5.13	3.84		3.26		
	J'' cm²/dyne(×10⁻⁹)	86.8		76.1	67.9		60.6	53.8	46.0	37.9		31.8	25.4	20.0	17.2	15.5		10.5	10.2	7.33		5.86		
14.8	J'	140.		110.	92.7		77.2	66.4	53.1	42.1		33.4	24.9	17.6	15.7	13.2	9.52	...	7.72	6.08		5.00		
	J'' cm²/dyne(×10⁻⁹)	96.4		88.8	81.9		74.8	71.6	61.3	49.6		41.0	33.0	26.3	23.6	20.9	15.4	...	14.3	10.9		8.21		
19.8	J'	176.		144.	124.		106.	89.5	71.0	57.1		46.9	35.7	26.1	21.8	18.8	...		10.9	8.92		7.04		
	J'' cm²/dyne(×10⁻⁹)	99.4		96.8	93.2		87.4	84.8	75.1	64.8		51.9	44.3	35.8	29.8	27.2	...		19.4	14.8		11.7		
25.0	J'	204.	184.	182.	156.		137.	119.	98.4	79.0		65.3	51.2	38.5	32.1	27.1	20.7		15.7	12.1		8.94		
	J'' cm²/dyne(×10⁻⁹)	102	93.6	107.	98.0		97.8	95.2	88.1	77.9		65.0	56.4	46.9	41.0	36.4	30.1		26.3	20.5		16.6		
30.2	J'		...		198.			154.	129.	109.		90.0	70.3	53.8	46.2	38.4	...		22.9	19.1		...		
	J'' cm²/dyne(×10⁻⁹)		...		103.			103.	98.6	84.2		79.6	70.8	61.2	51.4	47.6	...		35.5	26.8		...		
34.9	J'		...		227.		204.	187.	168.	142.		116.	92.1	73.7	59.7	51.9	44.2		...	26.3		...		
	J'' cm²/dyne(×10⁻⁹)		...		102.		99.0	102.	96.8	97.8		90.7	81.2	72.1	61.0	58.3	48.5		...	31.4		...		
40.0	J'		271.		252.		231.	217.	190.	167.		146.	117.	95.0	78.5	70.2	55.6		...	...		...		
	J'' cm²/dyne(×10⁻⁹)		76.7		89.0		94.0	98.3	93.0	94.1		97.4	94.3	81.1	73.7	69.2	52.7		...	...		...		
50.0	J'		307.		286.		278.	269.	251.	225.		199.	177.	141.	123.	113.	89.8		...	...		...		
	J'' cm²/dyne(×10⁻⁹)		55.7		64.2		79.2	81.9	91.2	92.1		95.2	101.	93.0	90.3	78.1	74.7		...	...		...		
59.8	J'		294.		295.		292.	282.	269.	260.		239.	221.	188.	...	...	...		...	...		...		
	J'' cm²/dyne(×10⁻⁹)		40.2		51.0		53.5	58.4	67.6	71.0		82.1	91.8	96.3	...	...	...		...	...		...		
80.0	J'		310.		319.		315.	...	302.	290.		282.	277.	257.	...	...	...		...	...		...		
	J'' cm²/dyne(×10⁻⁹)		32.3		34.9		36.8	...	40.2	43.5		50.0	60.3	71.3	...	...	...		...	...		...		
99.9	J'		302.		...		316.	314.	310.	303.		293.	295.	285.	...	...	...		...	...		...		
	J'' cm²/dyne(×10⁻⁹)		23.4		...		28.3	29.2	31.1	31.7		33.8	40.2	47.1	...	...	...		...	...		...		

[a]The frequencies of measurement were originally chosen to give nearly equal logarithmic spacing but were subsequently changed to provide a recognizable oscilloscope pattern when the importance of precise frequency settings became manifest (for example, 1400 was changed to 1500 cps, etc.). Three dots (···) signifies an error or an estimated uncertainty of more than 2 percent, as described previously in the main body of the paper. The measurements above 25.0°C were made with a source (later replaced) with so much distortion present at low frequencies that 30 cps measurements were not possible, although they could otherwise have been made.

The Calibration of Photographic Emulsions for Electron Diffraction Investigations

L. S. Bartell and L. O. Brockway
Department of Chemistry, University of Michigan, Ann Arbor, Michigan
(Received February 6, 1953)

A procedure for calibrating photographic emulsions has been devised which measures the deviation from linearity of the curve relating electron exposure to optical density. The experimental application of this procedure is found to give results consistent with a very simple model of the exposure process. It is also noted that the sensitivity of a photographic plate increases with increasing distance from the center of the plate. Measurements of this effect are given.

IMPROVEMENTS in experimental technique and analysis of data in the field of electron diffraction by gas molecules, most notably in the development of the rotating sector-microphotometer method, have made possible both more reliable and more detailed information about the location of atoms in molecules. Since electron diffraction patterns are recorded on photographic emulsions, it is essential for the most precise work that the response characteristics of the emulsions to exposure by fast electrons be accurately known. It is necessary to have a simple procedure for routine calibration of emulsions to allow for possible fluctuations in photographic materials.

A calibration method has been devised by Karle and Karle[1] for measuring the variation of optical density with exposure to electrons which relies upon the assumption that two gas diffraction plates taken under identical conditions except for exposure time have radial intensity distributions of exactly the same shape. A closely related method has been developed at the University of Michigan which depends upon the same assumption but studies the deviation from linearity of the optical density versus exposure curve. The main advantages of the Michigan method are that errors are more easily recognized and uncertainties averaged out, that the resulting calibration plot is as sensitive for small as for large optical densities, and that in many cases the conversion of density to exposure can be accomplished very simply analytically, much faster than by graphical procedures. This last advantage is especially important in converting the hundreds of individual values encountered in molecular structure determinations.

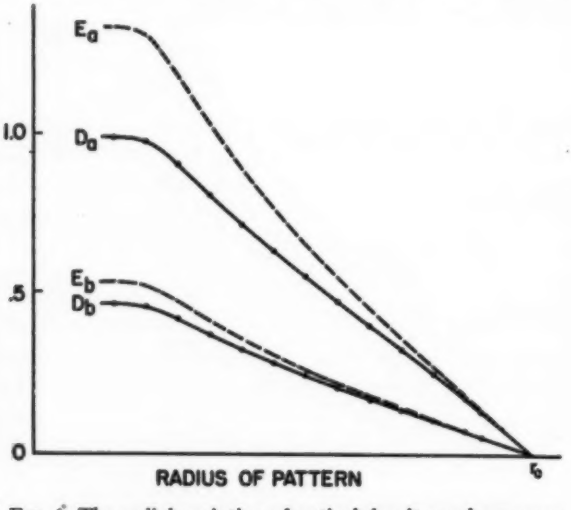


FIG. 1. The radial variation of optical density and exposure in a typical pair of calibration diffraction patterns.

The calibration procedure is as follows: Two gas diffraction patterns a and b are prepared so that the exposure of a is twice or somewhat more than twice that of b. The optical density D of each pattern is measured as a function of the radius, r, by a recording microphotometer while the pattern is being spun rapidly about its center of symmetry to average out irregularities in the emulsion.[2] It is desired to find the relation between D and E/D, where E is the exposure, the product of intensity and time. If $F(D)=E/D$, it is seen that the departure of $F(D)$ from constancy is a measure of the deviation from linearity of the D *vs* E curve. Both D and E are functions of the parameter r. The condition that the radial intensity distributions of a and b are identical in shape is that $E_a(r)=RE_b(r)$, where R is the ratio of the total exposures of the patterns. Figure 1 shows how D_a, D_b, E_a, and E_b vary with r in a typical pair of sectored calibration patterns taken with argon gas. At small D, $D\propto E$ approximately. For convenience the arbitrary proportionality constant is taken to be unity; so to the first approximation $F_0(D)=1$. Since this is a better approximation for b than for a, D_b being less than D_a, an improved approximation $F_1(D)$ can be found by noting that $E_a(r)=RE_b(r)\simeq F_0(D_b)RD_b$ and substituting this information into the defining equation for $F(D)$. The result is that $F_1(D_a)=F_0(D_b)RD_b(r)/D_a(r)$. This procedure for improving $F(D)$ can be carried out indefinitely, for $F_n(D_a)=F_{n-1}(D_b)RD_b(r)/D_a(r)$. In practice this series of approximations converges about as rapidly as the geometric series $\sum_n R^{-n}$. F_∞ can be found graphically or, if F_1 can be fit by a simple function, analytically. For instance, if $RD_b/D_a=1+cD_a$, it follows that

$$\frac{E}{D}\cong 1+\frac{cR}{(R-1)}D+\frac{2c^2}{R^2}\left(\frac{R^2+1}{R-1}\right)D^2.$$

[1] J. Karle and I. L. Karle, J. Chem. Phys. **18**, 957 (1950).

[2] (a) Degard, Pierard, and van der Grinten, Nature **136**, 142 (1935). (b) Karle, Hoober, and Karle, J. Chem. Phys. **15**, 765 (1947).

Figure 2 shows how the $F_n(D)$ were found to vary with D for the patterns in Fig. 1. It is to be emphasized that it is not necessary to know the ratio R in advance. R can be found by extrapolating the experimental function $D_a(r)/D_b(r)$ to $D=0$, but even this is not necessary because it is only the shape of the function $F(D)$ in the range of D covered in the diffraction investigation that is important. Knowledge of the shape of $F(D)$ for D greater than D_m requires knowledge of $D_a(r)/D_b(r)$ from D_a greater than about $D_m/2$.

In order to determine $F(D)$ completely it is necessary that the density of both diffraction plates vary over a large range. On the other hand, it is desirable from photometric considerations regarding accuracy that the density of a pattern to be measured vary between as narrow limits as is possible. A carefully shaped rotating metal mask known as a rotating sector is therefore employed to modify in a known way the intensity of diffracted electrons reaching the photographic plate. The intensity of an unsectored electron diffraction pattern falls off roughly as $r^{-3.5}$, which is too rapid a drop to be suitable for calibration purposes, and the r^3 sector used at Michigan to record molecular diffraction patterns gives too flat a pattern. For these reasons a special calibration sector with an angular opening proportional to $r^3(r_0-r)$ has been constructed that can be easily placed over the regular sector. The diffraction patterns obtained with this sector have the convenient form illustrated in Fig. 1.

It is assumed in the procedure described by Karle and Karle[1] that $F(D)$ is constant from densities of zero to about 0.25 for contrast lantern slides. The method described above does not require this assumption, and the results obtained by it indicate that such an assumption will lead to an error of several percent in the calibration curve at optical densities less than 0.5. For contrast lantern slides it appears that $F(D)=1+cD$, with $c\sim 0.3$, for optical densities up to 0.5. It is interesting that this is just the form that would be expected at small D for the following highly simplified model of the exposure process: (A) The photographic emulsion contains grain domains which are completely blackened by the impact of one fast electron and which are unaffected by further impact. (B) The impact of an incident electron affects only one domain. The optical density, which is proportional to the number of blackened grains per unit area, increases less rapidly than the number of electron impacts per area because of multiple impacts upon some of the grains. The relaxation of either condition A or B can result in a decreasing slope for $F(D)$ as D approaches zero. These considerations suggest the possibility of applying this type of investigation to the study of the mechanism of exposure processes.

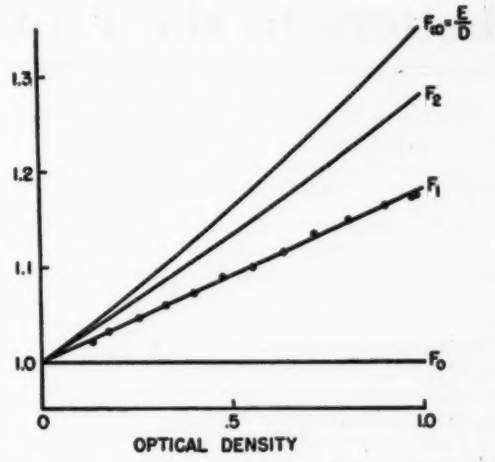


FIG. 2. The variation of $F_n(D)$ with optical density, The circles represent experimental values of RD_b/D_a computed from data plotted in Fig. 1.

Another property of photographic plates that must be taken into account in intensity measurements is that the sensitivity of the emulsion varies with position on the plate. The edges of the plates are more sensitive than the centers. It is well known that mechanical disturbances can render an emulsion more sensitive; cutting and packing the plates with spacers at the edges may account for the variation. In order to get a measure of the magnitude of the effect, some Kodak $3\frac{1}{4}$ in. $\times$ 4 in. lantern slide plates were given uniform photon exposures varying from 0.2 to 1 in optical density and spun rapidly about their centers while being microphotometered on a Leeds and Northrup recording microphotometer. The traces were remarkably smooth but not flat. The variation of sensitivity with the radius r could be expressed by the same relationship for all of the exposures to within several parts per thousand for radii smaller than $3\frac{1}{2}$ cm. If this were not true, the calibration method just described would not be valid. The sensitivity was taken to be the ratio of the apparent exposure $E'(r)$ at r to the apparent exposure at $r=0$, where $E'(r)=D(r)F(D)$. The variation was approximately $E'(r)/E'(0)=1+0.004r^2$ with r in cm. For radii greater than $3\frac{1}{2}$ cm, $E'(r)/E'(0)$ increased more rapidly than $(1+0.004r^2)$, and the experimental points started to scatter.

Although this sensitivity variation is important in the study of scattering of electrons by atoms (to be published), fortunately it does not affect electron diffraction investigations of molecular structure, since such investigations are concerned with the ratio of molecular to atomic intensities rather than with direct intensities.

Letters to the Editor

Titanium Replica for Electron Microscopy

SHIGETO YAMAGUCHI AND TADAYUKI NAKAYAMA
Scientific Research Institute, Ltd., Hongo, Tokyo, Japan
(Received February 9, 1953)

A SMALL piece of pure spongy titanium coiled in tungsten wire (diameter: 0.2 mm) heated by electric current flowing through it was melted and evaporated in vacuum to form a thin film of the metal on a glass plate. When the "mirror" thus obtained was dipped in water at about 15°C, the thin film (thickness: 200–500A) was readily freed from the glass. It should here be remarked that the metallic film deposited on the cooler domains of a glass plate can be more easily freed than the deposit on the warmer domains nearer to the hot tungsten filament.

The metal film prepared in this way was studied by means of electron microscopy and microdiffraction [see Figs. 1(a) and (b)]. These two figures show that there is no granular coarseness in the metal film and that the film is composed mainly of titanium metal along with a small amount of tungsten and titanium dioxide (rutile).

The metal film of Fig. 1 was oxidized in air at about 300°C for about 10 hours. The oxide film thus obtained is shown in Figs. 2(a) and (b). It is seen from Figs. 2(a) and (b) that oxides composing the film are not granular, although some wrinkles have appeared here, and that the film is now composed mainly of titanium dioxide (rutile).

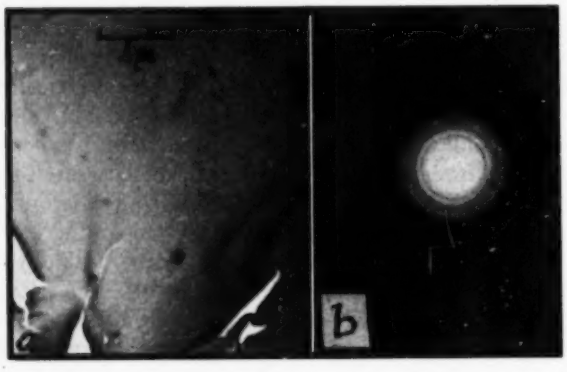


FIG. 1. (a) Titanium film formed on a glass plate and removed. The film is not granular. (b) A microdiffraction pattern obtained from the film of Fig. 1(a). The pattern is composed of continuous rings from titanium, tungsten, and rutile.

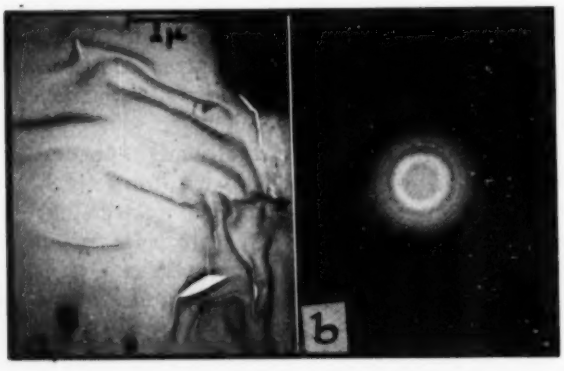


FIG. 2. (a) The micrograph of the oxide film obtained by thermally oxidizing the film of Fig. 1. (b) A diffraction pattern of Fig. 2(a). This pattern verifies the existence of rutile, whose crystallites are not granular.

FIG. 3. (a) The micrograph of the titanium film formed on a glass plate scratched with emery paper. This micrograph with straight polish marks implies a possibility of taking replica of optical grating with a titanium film. (b) Traces of scratch formed on a glass surface by polishing.

Replica for electron microscopy.—Information obtained in the preliminary examination infers that a titanium film prepared by evaporation in vacuum might be able to reprint a glass surface. As a matter of fact, the titanium film formed on a glass plate and isolated therefrom gave Figs. 3(a) and (b). This glass had been scratched with a comparatively coarse emery paper (05). Figure 3(a) infers that titanium films can be useful for making replicas of optical gratings for magnification measurement in electron microscopy.

Since the titanium film does not yield oxides subject to granular coarsening even after thermal oxidation, it may be a stable and heat-resisting material for replicas in electron microscopy.

It is very difficult to divorce the thin film of titanium formed on an etched surface of stainless steel. This implies that the titanium film prepared by evaporation in vacuum is not suitable for making replicas of metal surfaces.

Effect of Internal Strains on Linear Expansion, X-Ray Lattice Constant, and Density of Crystals

CHESTER R. BERRY
Research Laboratories, Eastman Kodak Company, Rochester, New York
(Received February 16, 1953)

IN recent publications Miller and Russell[1,2] discuss the effects in crystals of internal strains produced by defects on the linear expansion, x-ray lattice constant, and density. The results of their theoretical analysis show, for the type of distortion which they assume, that it is not necessary for the lattice expansion to equal the macroscopic expansion produced by the defects. In fact, they estimate that the lattice-constant change, if it may be measured by the change in angle of the maximum of the diffraction line, is just about twice as great as the macroscopic change, i.e., $\Delta a/a/\Delta l/l=\alpha=2.0$.

The purpose of this note is to indicate that measurements reported in the literature allow an evaluation of α to be made directly for some cases and that the value is approximately 1.0

rather than 2.0. One such case is for the ionic crystals KBr and KCl in which solid solution occurs. The lattice constant as a function of composition is given by Oberlies.[3] For the addition of x mole percent of KBr to KCl the lattice-constant change is $\Delta a/a=+0.00051x$. Changes in the volume and length of such crystals may be obtained from the results of density as a function of composition as given by Tammann and Krings.[4] The density values lie very close to a straight line when plotted against composition in mole percent. Using the straight line, $\Delta l/l=+0.00073x$; then the value of $\alpha=0.70$. For additions of KCl to KBr the values are $\Delta a/a=-0.00048y$ and $\Delta l/l=-0.00043y$ and $\alpha=1.12$. If vacancies or voids were produced in forming the solid solutions, the values of $\Delta l/l$ cited above would be somewhat too large without corresponding increases in $\Delta a/a$. This would tend to make the correct values of α more nearly equal to 1.0 in both cases.

Similar computations may be made for a metallic crystal containing a known amount of substitutional impurity. Such a system involving aluminum added to silver is described by Foote and Jette.[5] They find accurate straight-line relationships of both density and lattice constant as a function of aluminum content up to 15 atom percent. It is found that $\Delta a/a=-0.000308x$, $\Delta l/l=-0.000280x$, and $\alpha=1.10$.

Measurements of density and lattice constant of silver with zinc impurity have been made.[6,7] Special care was taken to assure the uniformity and soundness of these alloys. For small additions of zinc the following values were found: $\Delta a/a=-0.000446x$, $\Delta l/l=-0.000453x$, and $\alpha=0.99$.

In view of these results, it appears that the usual assumption of $\alpha=1.0$ may be generally applicable.

[1] P. H. Miller, Jr., and B. R. Russell, Phys. Rev. 86, 650 (1952).
[2] P. H. Miller, Jr., and B. R. Russell, J. Appl. Phys. 23, 1163 (1952).
[3] F. Oberlies, Ann. Physik 87, 238 (1928).
[4] G. Tammann and W. Krings, Z. anorg. Chem. 130-132, 242 (1923).
[5] F. Foote and E. R. Jette, Trans. Am. Inst. Mining and Met. Engrs. 143, 151 (1941).
[6] D. Stockdale, J. Inst. Metals 66, 287 (1940).
[7] Lipson, Petch, and Stockdale, J. Inst. Metals 67, 79 (1941).

Reflection Coefficients for Certain Rough Surfaces*

V. TWERSKY
Mathematics Research Group, New York University, New York, New York

(Received February 23, 1953)

IN a recent paper[1] we considered the scattered reflection of electromagnetic waves from perfectly conducting surfaces composed of either semicylindrical or hemispherical bosses on a smooth plane. In this letter we obtain reflection coefficients for these surfaces by equating the scattered energy resulting from unit area of incident wave with the energy "missing" from unit area of the wave reflected at the specular angle. We then indicate how the identical results may be obtained analytically. We assume single scattering but will indicate the range of validity of the results on the basis of our multiple scattering treatment of the cylindrical case.[2]

The incident wave is specified by $\exp[-ikr\cos(\theta-\alpha)-i\omega t]$, where α and θ, the angles of incidence and observation, are measured from the normal of the plane. The incident polarization amplitudes are taken to be unity perpendicular and parallel to the plane of incidence; the corresponding scattered fields are indicated by the subscripts $\perp$ and $\|$. We first consider the two-dimensional cylindrical problem and then the spherical problem. The acoustic case is also treated.

The far-field form of the wave scattered by a semicylindrical boss of radius $a\ll k^{-1}=\lambda/2\pi$ is specified by[1]

$$\begin{Bmatrix} E_\perp \\ E_\| \end{Bmatrix}=\left(\frac{\pi i}{2kr}\right)^{\frac{1}{2}}(ka)^2 e^{ikr}\begin{Bmatrix} 2\cos\alpha\cos\theta \\ -1-2\sin\alpha\sin\theta \end{Bmatrix}. \tag{1}$$

The number of scatterers excited by unit area of incident wave is $\rho\sec\alpha$, where ρ is the number of bosses in unit length of surface. Hence the normalized energy flux single scattered from unit are of incident wave, or the "scattering cross section of the surface per unit area of incident wave," is (for either component)

$$\sigma=q\rho\sec\alpha=\rho\sec\alpha\int_{-\pi/2}^{\pi/2}|E|^2 r\,d\theta,\quad r\to\infty, \tag{2}$$

where q is the scattering cross section per unit length of a single boss. For $\rho/k\ll 1$ we take σ to be the energy "missing" from unit area of the wave reflected at the specular angle $\theta=-\alpha$. Thus we write the reflection coefficients as

$$R=1-\sigma\approx e^{-\sigma}, \tag{3}$$

where it is understood that $\sigma\ll 1$; specifically we have

$$\begin{Bmatrix} R \\ R_\| \end{Bmatrix}=1-\frac{\rho}{k}\pi^2(ka)^4\begin{Bmatrix} \cos\alpha \\ (1+2\sin^2\alpha)/2\cos\alpha \end{Bmatrix}. \tag{4}$$

$R_\|$ of Eq. (4) breaks down near grazing incidence ($\alpha\to\pi/2$) as a consequence of multiple scattering.[3]

For the analogous problem of the hemispherical boss we have[1]

$$\begin{Bmatrix} E_\perp \\ E_\| \end{Bmatrix}=\frac{k^3a^3}{r}e^{ikr}\begin{Bmatrix} \cos\alpha & (\cos\boldsymbol{\phi}\cos\theta\cos\varphi+\boldsymbol{\theta}\sin\varphi) \\ \boldsymbol{\phi}\cos\theta\sin\varphi-\boldsymbol{\theta}(2\sin\theta\sin\alpha+\cos\varphi) \end{Bmatrix}, \tag{5}$$

where $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$ are unit vectors; θ is the polar angle and φ is measured from the plane of incidence. The scattering cross section of the surface per unit area of incident wave is

$$\sigma=q\rho\sec\alpha=\rho\sec\alpha\int_0^{2\pi}d\varphi\int^{\pi/2}|E|^2 r^2\sin\theta\,d\theta, \tag{2'}$$

where q is the scattering cross section of a single boss and ρ is the number of scatterers in unit area of surface. The reflection coefficients are thus

$$\begin{Bmatrix} R_\perp \\ R_\| \end{Bmatrix}=1-\frac{\rho 4\pi(ka)^6}{k^3 3}\begin{Bmatrix} \cos\alpha \\ (1+4\sin^2\alpha)/\cos\alpha \end{Bmatrix}, \tag{6}$$

where $\rho/k^3\ll 1$. $R_\|$ breaks down as $\alpha\to\pi/2$.

We now consider the corresponding acoustic problems. For the cylindrical case $E_\perp$ and $E_\|$ of Eq. (1) are the velocity potentials for a free and rigid surface respectively; consequently, the results of Eq. (4) apply. For the spherical case this correspondence does not hold; instead the scattered velocity potentials for the single bos are given by[1]

$$\begin{Bmatrix} \psi_f \\ \psi_r \end{Bmatrix}=\frac{k^3a^3}{r}e^{ikr}\begin{Bmatrix} 2\cos\alpha\cos\theta \\ (-2-3\sin\alpha\sin\theta\cos\varphi)/3 \end{Bmatrix}. \tag{7}$$

Consequently,

$$\begin{Bmatrix} R_f \\ R_r \end{Bmatrix}=1-\frac{\rho 8\pi(ka)^6}{k^3 3}\begin{Bmatrix} \cos\alpha \\ [1+(\frac{3}{4})\sin^2\alpha]/3\cos\alpha \end{Bmatrix}, \tag{8}$$

where $\rho/k^3\ll 1$; R_r breaks down as $\alpha\to\pi/2$.

The intuitive procedure we used above can be justified and generalized analytically for a large class of nonabsorbing surfaces. Consider the two-dimensional problem of a wave incident on an arbitrary configuration of identical (not necessarily circular) bosses on an infinite plane. We calculate the total time-average Poynting flux for a particular surface (i.e., a particular configuration) and average the flux over the ensemble of surfaces. If the distribution is uniform, then the normalized average reflected flux to lowest order in ρ (actually $\rho/k\cos\alpha$) is

$$\langle I\rangle=\left\{1+2\rho Re\left[E_p^*\int E(y')dy'\right]\right\}n+\rho\int|E(y')|^2 r_0'dy', \tag{9}$$

where ρ is the average number of elements in unit length, $E(y')$ is the single-scattered wave of a boss located at $x=0$, $y=y'$, and E_p is the wave reflected from the plane in the absence of the bosses.[4] $n=i\cos\alpha-j\sin\alpha$ is a unit vector in the specular direction, and $r_0'=i\cos\theta'+j\sin\theta'$ is a unit vector from $x=0, y=y'$ to the observation point x, y. We now show that the interference term of Eq. (9) equals σ of Eq. (2) for a boss of arbitrary cross section.

Since the component of the reflected flux normal to the surface must equal $\cos\alpha$ for a nonabsorbing surface, we obtain from Eq. (9) the relation

$$-2\cos\alpha R\epsilon\left[E_p^*\int E(y')dy'\right]=\int|E(y')|^2\cos\theta'dy', \quad (10)$$

where $\cos\theta'=x/r'=x[x^2+(y-y')^2]^{-\frac{1}{2}}$. We now make use of $d\theta'=-\cos\theta'dy'/r'$ and change to the variable θ' on the right side of Eq. (10); then, if the limits in the integrals of Eq. (10) are infinite, we obtain

$$-2\cos\alpha R\epsilon\left[E_p^*\int_{-\infty}^{\infty}E(y')dy'\right]=\int_{-\pi/2}^{\pi/2}|E(\theta')|^2r'd\theta'=q. \quad (11)$$

Thus the terms arising from interference of the specular and scattered waves in Eq. (9) are equal to $-\rho\sec\alpha$ times the scattering cross section of a single boss, i.e., to σ. [We have also proved theorem Eq. (11) directly for a single boss without recourse to probability considerations.]

For the special case of a circular boss we employ the complete form of $E(y')$ and find that either the left or right side of Eq. (11) equals

$$\begin{Bmatrix}q_\perp\\q_\parallel\end{Bmatrix}=\frac{8}{k}\sum_{n=0}^{\infty}\begin{Bmatrix}|A_n|^2\sin^2n(\alpha+\pi/2)\\|B_n|^2\cos^2n(\alpha+\pi/2)\end{Bmatrix}, \quad |A_n|^2=\frac{J_n^2}{J_n^2+N_n^2},$$
$$|B_n|^2=\frac{J_n'^2}{J_n'^2+N_n'^2}, \quad (12)$$

where the argument of the cylindrical functions is ka and where the prime indicates differentiation with respect to ka. The above may be readily extended to the three-dimensional case.

We now show that Eq. (3) is a consequence of Eq. (9) and Eq. (11). We may apply $\langle I\rangle$ of Eq. (9) to a practical set-up, provided that $\rho/k\cos\alpha\ll1$ and provided that the width and height of the bosses are small compared to ρ^{-1}; the range of integration is to be taken as the illuminated region. If the incident beam width, say d, is very much larger than all wavelengths involved, then we may, in general, use infinite limits in the n component: because of the exponent of $E(y')$ the main contribution of the integral arises from the immediate vicinity of the specular or stationary point and is consequently practically independent of the width of the illuminated region.[5] Hence for the n component we may use Eq. (11) and the definition $\sigma=q\rho\sec\alpha$ to write the leading term of Eq. (9) as $(1-\sigma)n$. For the r_0' component (the "incoherent scattering"), however, the integrand is slowly varying, and the integration should be taken only over the illuminated region. If this region is small compared to the distance r of the observation point from its center (i.e., if $d\sec\alpha/r\ll1$), then we may write this term as $\rho d\sec\alpha|E(r,\theta)|^2r_0$. The function $|E(r,\theta)|^2$ is the normalized scattered energy flux of a boss at the center of the illuminated region, r_0 is the unit vector from this boss to the observation point, and $\rho d\sec\alpha$ is the number of illuminated bosses. (For small circular bosses $|E|^2$ may be obtained directly from Eqs. (1), (5), and (7).]

Hence we may write

$$\langle I\rangle\approx(1-\sigma)n+\rho d\sec\alpha|E|^2r_0. \quad (13)$$

This expression is to be used as follows: at the specular angle we ignore the term in $|E|^2$ (i.e., the energy scattered *into* the reflected beam), since it is only of the order of d/r times σ; thus we have

$$\langle I\rangle\cdot n\approx1-\sigma=R, \quad \theta=-\alpha. \quad (14)$$

We have therefore justified the use of Eq. (3), provided that $\rho/k\ll1$ and $\alpha\not\to\pi/2$. On the other hand, if θ is not near $-\alpha$, we ignore the first term of Eq. (13) and use

$$\langle I\rangle\cdot r_0\approx\rho d\sec\alpha|E|^2, \quad (15)$$

i.e., outside of the reflected beam we consider only the incoherent scattering.

We note that the results can be extended to the case of absorbing bosses by replacing σ by $\sigma+\sigma_a=(q+q_a)\rho\sec\alpha$, where q_a is the absorption cross section. However, it is simpler for this case to obtain R directly from the n component of Eq. (9). Thus if we write $E(y')=(2|i\pi kr)^{\frac{1}{2}}e^{ik(r'-y'\sin\alpha)}f(\theta'\alpha)$, where $f(\theta',\alpha)$, is a slowly varying function of y' compared to the exponent, then[2] $\int E(y')dy'\approx E_p(2/k\cos\alpha)\times f(-\alpha,\alpha)$. Consequently, we have

$$\langle I\rangle\cdot n\approx1+(4\rho/k\cos\alpha)R\epsilon[f(-\alpha,\alpha)]=R, \quad (16)$$

where $f(-\alpha,\alpha)$ is the value of the scattering amplitude at the specular angle. (Most generally we express $f(\theta,\alpha)$ as an integral over the surface of the boss.) It can be shown that R of Eq. (16) is identically $1-(\sigma+\sigma_a)$. The leading term of R of Eq. (16) involving ka (or an equivalent parameter) is of the lowest order of magnitude found in E, e.g., $R_{11}=1-\mathcal{O}(ka)^2$ for the circular cylindrical case, etc. [Compare with Eq. (14).]

We have assumed in the above that the incident beam is perfectly collimated and also that the pick-up is highly directional. In practice, one would take into account the radiation characteristics of both the source and pick-up. Note also that since our expressions are averages they should be compared with the average of measurements over a large number of appropriate surfaces.

* This work was performed at Washington Square College of Arts and Science, New York University and was supported in part by Contract No. AF-19(122)-42 with the U. S. Air Force through sponsorship of the Geophysics Research Directorate of the Air Force Cambridge Research Center, Air Research and Development Command.

[1] V. Twersky, J. Appl. Phys. 22, 825 (1951); see also J. Acoust. Soc. Am. 22, 539 (1950); 23, 336 (1951) for the acoustic case.

[2] V. Twersky, *Multiple Scattering of Waves by Planar Random Distributions of Parallel Cylinders and Bosses.* Mathematics Research Group, New York University (to be published).

[3] We found (see reference 2) that both $R_\perp$ and $R_{11}\to1$ as $\alpha\to\pi/2$; furthermore, the total intensity $\to0$ as $-\theta=\alpha\to\pi/2$ regardless of polarization. We used the probability distribution functions of liquid state theory to define the ensemble of surfaces, the "randomness" of the distrubition being specified by the ratio of the average separation of two scatterers to their minimum separation. We found the behavior at $-\theta=\alpha\to\pi/2$ to be independent of this ratio (and also of the dielectric properties of the bosses and of the value of ka), provided that the minimum separation was large compared to λ and a.

[4] Note that for the circular boss, the n component of Eq. (9) yields the results of Eq. (4), provided that terms to the order of $(ka)^4$ are retained in $E(y')$ and provided that the limits on the integral are infinite. We did not get the present result in reference 1 since we kept only the $(ka)^2$ terms of $E(y')$, i.e., the results of reference 1 are, in general, correct only to the order of $(ka)^2$ or $(ka)^3$.

[5] If this condition is not fulfilled, then we may use the radial energy fluxes of reference 1 for the finite distributions to the order of $(ka)^2$ or $(ka)^3$. For this case the flux may show a minimum at the specular angle,[1] i.e., the width of the "interference beam" is narrower than the width of the specular beam represented by 1 in Eq. (9), and consequently, the value at $\theta=-\alpha$ may be smaller than the values in the immediate vicinity of $-\alpha$.

Growth of Cadmium Sulfide Crystals

M. E. Bishop and S. H. Liebson
Naval Research Laboratory, Washington, D. C.
(Received February 23, 1953)

CADMIUM sulfide crystals are generally grown by two methods, one conceived by Frerichs[1] and the other recently described by Czyzak *et al.*[2]

Frerichs' method of growing cadmium sulfide crystals consists in the passage of H_2, essentially a vehicular gas, over molten cadmium, the mixture combining at an orifice with a concentric gas flow of H_2S to react to form cadmium sulfide crystals.

The method described by Czyzak *et al.* produces thicker crystals and involves their growth from powdered cadmium sulfide and H_2S in a sealed tube heated to 1000°C for several days and then slowly cooled. We have modified the Frerichs' technique to provide more effective conversion of the cadmium and H_2S into cadmium sulfide crystals (see Fig. 1).

The first modification consisted in the replacement of H_2 by argon as a vehicular gas. The vehicular gas flow is adjusted to low velocities of the order of 1 cm per second to prevent back diffusion of H_2S of sulfur from the reaction area. The cadmium was maintained at 650°C, which for the physical arrangement used delivered cadmium to the reaction area at the rate of approximately 6 grams per hour. The H_2S flow was adjusted until no cadmium appeared in the exhaust of the growth tube to provide a more effective utilization of the cadmium charge. For concentric

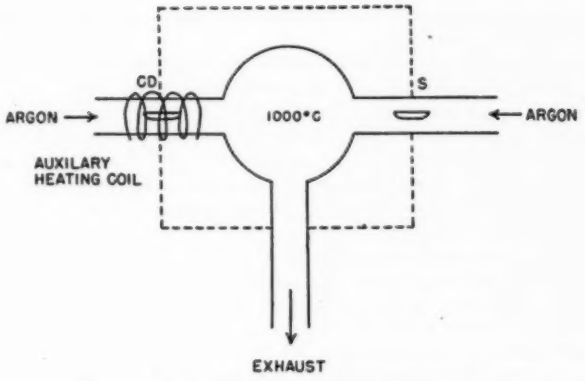


FIG. 1. Crystal growing apparatus.

quartz tubes of greater than 2.5-cm diameter is was found advisable to use baffles at both sides of the crystallization area to minimize concentric currents.

The best crystal growths occurred when a dense yellow fog of sulfur appeared in the exhaust tube, lending support to the belief that the improvement of crystal growth obtained by the use of argon instead of H_2 is due to the greater density of sulfur available in the reaction area.

Another method of growing cadmium sulfide crystals which we have successfully used is illustrated in the accompanying figure. Argon is passed over molten cadmium sulfide from one side and molten sulfur from the other to combine in a quartz balloon flask, with exhaust going out a third tube. The reaction area is maintained at close to 1000°C, the crystals forming inside the flask.

The use of H_2S in the right arm to replace the sulfur and argon, also proved equally effective.

This arrangement suffers from the lack of adaptability of the growing apparatus to commercial furnaces but has the advantage that the cadmium temperature can be maintained more constant by auxiliary heating coils around the quartz tube and is difficult in the Frerichs method because of the necessity for concentric tubes to surround the cadmium boat at this point.

Crystals grown according to this method ranged in size up to several square centimeters in area, one millimeter thick, and did not differ in shape, color, fluorescence, or electrical properties from crystals grown by the method described by Frerichs or the modifications noted herein.

The purity of all gases used was 99.8 percent or better; the cadmium was spectroscopically analyzed as 99.98 percent pure, with less than 0.002 percent Fe, 0.010 percent Pb, and 0.005 percent Cu.

[1] R. Frerichs, Phys. Rev. **82**, 594 (1947).
[2] Czyzak, Craig, McCain, and Reynolds, J. Appl. Phys. **23**, 928 (1947).

Comments on Louis A. Pipes' Letter[1]

H. S. KIRSCHBAUM
Ohio State University, Columbus, Ohio
(Received January 8, 1953)

PROFESSOR Pipes in his comments upon my letter of August 4, 1952 dismissed the experimental evidence showing that i_1 does not equal i_2 and that (i_1+i_2) does not contain a sustained unidirectional component by implying that in some way this amplifier is special. That is, that it does not meet "the symmetry of the configuration of the circuit on which the analysis is based." That the experimental configuration is symmetrical is borne out by the similarity which exists between i_1 and i_2; that the amplifier is special is perhaps true to the extent that R_w/R_L is considerably less than would be found in practice. If practical leakage fluxes were responsible for such extreme discrepancies between experimental and analytical results, it is apparent that any analysis which does not consider leakage flux is of doubtful value. The author in his reply to my letter still has not dispelled the air of confusion surrounding the two aforementioned points. I would like to take each of these points up in turn.

The author concludes that i_1 equals i_2 by using Eqs. (2–2) and (2–3) without any assumption of "symmetry" to start with. The conclusion is based upon a process which makes no use of the manner in which ϕ_a and ϕ_b depend upon i_1 and i_2. This in itself should be sufficient to show the fallacy of the analytical procedure which could be applied in the same formal manner to a linear circuit in which $N\phi_a=L_a i_1$ and $N\phi_b=-L_b i_2$. If R_w were zero, the only valid conclusion which could be drawn in this linear case would be that $L_a(di_1/dt)$ equals $L_b(di_2/dt)$, which certainly does not mean $i_1=i_2$. The new analytic formulation used by the author in his letter does not dispel the fallacies of his original procedure. It is stated in Eqs. (6) and (7) of the letter,

$$i_1=\frac{(1+k)F_1-F_2}{2k+k^2}, \tag{6}$$

$$i_2=\frac{(1+k)F_2-F_1}{2k+k^2}. \tag{7}$$

The author correctly recognizes that as k approaches zero, F_1 and F_2 must approach each other in order to satisfy the condition that i_1 and i_2 be finite. A common value F is substituted in Eqs. (6) and (7) for F_1 and F_2, and then i_1 and i_2 are solved for, letting k approach zero. This procedure yields $i_1=i_2=F/2$. That this is incorrect can be seen as follows. Equations (6) and (7) can be rewritten as

$$i_1=\frac{(F_1-F_2)/k+F_1}{(2+k)} \tag{6a}$$

$$i_2=\frac{(F_2-F_1)/k+F_2}{(2+k)}. \tag{7a}$$

Now if we let k approach zero, i_1 and i_2 will approach each other if and only if $(F_1-F_2)/k$ approaches zero. This is an indeterminate form (being of the form zero over zero). There is nothing in the physics of the problem to warrant throwing this term away as Professor Pipes has done. Consequently, the conclusion that i_1 equals i_2 is unwarranted.

With regard to the presence of direct current in the load, the quotation of Helmholtz's "Theory of Combination Tones" hardly seems relevant. Of course, sustained direct components are found in many nonlinear instances (a most trivial example is offered by the direct component of current flowing in the secondary of an ac transformer whenever a nonlinear resistor of the rectifier type is in series with a linear load resistance). But this fact does not exclude the possibility that there may be many other nonlinear situations in which rectification does not occur. We can see that there is no direct current in the load by considering Eqs. (2–2) and (2–3) of the original paper. Let us define steady-state operation as being characterized by the fact that each cycle of operation is exactly like the preceding one. With the amplifier operating under steady-state conditions, if Eqs. (2–2) and (2–3) are added and integrated over a complete cycle, it is immediately verifiable that the integral of (i_1+i_2) over one cycle is zero—hence no direct current can be found in the solution.

[1] Louis A. Pipes, J. Appl. Phys. **23**, 1279 (1952).

Correction to "A Mathematical Analysis of Parallel-Connected Magnetic Amplifiers with Resistive Loads"

LOUIS A. PIPES
University of California, Los Angeles, California
(Received March 3, 1953)

PROFESSOR Kirschbaum and Mr. A. C. Sim of Standard Telecommunication Laboratories Limited, Enfield, Middle-

sex, England, have clearly pointed out that the argument on which the analysis of the paper, "A Mathematical Analysis of Parallel-Connected Magnetic Amplifiers with Resistive Loads,"[1] is based is fallacious. The correct evaluation of the indeterminate form encountered in the analysis indicates that the difference $(i_1 - i_4)$ is an *arbitrary* quantity and not necessarily zero as was taken in the paper.

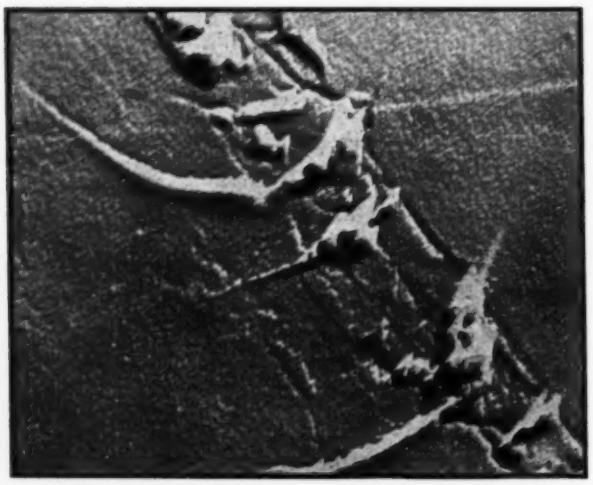
The author wishes to thank Professor Kirschbaum and Mr. Sim for pointing out this error. He deeply regrets the inconvenience and trouble this matter has caused.

[1] Louis A. Pipes, J. Appl. Phys. 23, 625 (1952).

A Method of Preparing Replicas for Electron Microscopic Examinations

W. S. Smith and W. G. Kirchgessner
Bausch and Lomb Optical Company, Rochester, New York
(Received January 23, 1953)

IN the present paper a modification of the Formvar process[1] for making replicas of specimens for electron microscopic examination is described. The proposed method is based on a rapid two-step process which results in improved replicas for the examination of surface characteristics.



(a)

(b)

FIG. 1. Electron micrographs of a scratched glass surface at 15 000× magnification (areas identical). (a) Prepared from Formvar-chromium replica. (b) Prepared from chromium-silica replica.

In preparing replicas of solid surfaces the specimens are first coated in the usual manner with a layer of Formvar which is then stripped onto specimen screens and placed under vacuum. Following the evaporation of a shadowing metal[2] (chromium), a silica film is deposited on the replicas from a filament mounted directly above them. This material has been used previously in the preparation of replicas and was first reported in a paper by Heidenreich and Peck.[3]

During the evaporation of both the chromium and the silica the pressure in the evaporating unit is maintained below 10^{-4}mm Hg. Silica in the form of quartz splinters placed in a coiled tungsten filament 4 inches above the replicas is evaporated at 15 amperes for 10 minutes. The resulting silica layer is thick enough to resist tearing and thin enough for the electron beam to penetrate. As measured by a Michelson interferometer this film was found to be approximately 100A in thickness and is simply controlled by timing the evaporation at a given amperage.

After evaporation the silica coated replicas are removed from the evaporating chamber and are immersed in a bath of ethylene dichloride for 30 seconds. The Formvar layer is dissolved in this bath, leaving the shadowing metal and the substituted silica layer adhering to the supporting screen. Thus the detail in the surface of the Formvar layer which was in contact with the specimen is preserved in the shadowed silica film, while any artifacts[4-6] associated with the use of Formvar are eliminated.

The method herein presented lends itself readily to a reliable and rapid examination of surface contours. This conclusion is supported by the examination of the photogrpahs in Fig. 1, which show that an appreciable gain in photographic detail may be realized through the use of the proposed technique.

[1] V. J. Schaefer, and D. Harker, J. Appl. Phys. 13, 427 (1942).
[2] R. C. Williams and R. W. G. Wyckoff, J. Appl. Phys. 15, 712 (1944).
[3] R. D. Heidenreich, and V. G. Peck, J. Appl. Phys. 14, 23 (1943).
[4] Deacon, Ellis, Cross, and Sennett, J. Appl. Phys. 19, 704 (1948).
[5] J. Trotter, Nature 164, 227 (1949).
[6] I. W. Fischbein, J. Appl. Phys. 21, 1199 (1950).

Fourier Representation of Buerger's Image-Seeking Minimum Function

William J. Taylor
McPherson Chemical Laboratory, Ohio State University, Columbus, Ohio
(Received March 2, 1953)

BUERGER[1] has shown that Patterson syntheses in which the peaks are sufficiently well resolved may be solved for the atomic positions (and, therefore, the *phases* of the structure factors) by image-seeking methods. These methods depend on the definition of a suitable function of the Patterson values at the vertices of an *image polygon* such that, as the polygon is translated over the unit cell, the function assumes a peak value when the polygon is in an image position and has a low value otherwise (except possibly for accidental coincidences which become less probable as the number of vertices in the polygon is increased). The most useful type of image-seeking function, at least in the early stages of an analysis, seems to be the line image function, which depends on the two Patterson values at the end of a line or vector. This vector is chosen to correspond with a principal interatomic vector, as judged from a preliminary study of the Patterson synthesis or its projections, and a knowledge of the symmetry of the crystal. In the ideal case of perfectly sharp Patterson peaks (i.e., points) the use of the line image yields the atomic positions directly (plus a mirror image for noncentrosymmetric structures). The only theoretical limitation on the method is that it will not separate so-called homometric structures.[2]

The most effective image-seeking function introduced by Buerger[1c,d] is the *minimum function*, defined as follows (for a

centrosymmetric structure):

$$M_{uvw}(x,y,z)=\text{Min}\{P(x+\tfrac{1}{2}u,\, y+\tfrac{1}{2}v,\, z+\tfrac{1}{2}w),\\ P(x-\tfrac{1}{2}u,\, y-\tfrac{1}{2}v,\, z-\tfrac{1}{2}w)\}. \quad (1)$$

The right-hand side of Eq. (1) represents the minimum of the two Patterson values indicated, if they are unequal, or their common value, if equal. The values of the Patterson function and therefore of $M_{uvw}(x,y,z)$ are inherently positive (or zero). The present note is concerned with a Fourier series representation of the function $M_{uvw}(x,y,z)$ of Eq. (1). Buerger, in his application of the minimum function, has used a visual shifted-Patterson technique.[1d] It is possible to represent $M_{uvw}(x,y,z)$ conveniently as the difference of two Fourier series in the following way. Define the *sum* and *difference functions*,

$$S_{uvw}(x,y,z)=\tfrac{1}{2}[P(x+\tfrac{1}{2}u,\, y+\tfrac{1}{2}v,\, z+\tfrac{1}{2}w)\\ +P(x-\tfrac{1}{2}u,\, y-\tfrac{1}{2}v,\, z-\tfrac{1}{2}w)], \quad (2)$$

$$D_{uvw}(x,y,z)=\tfrac{1}{2}[P(x+\tfrac{1}{2}u,\, y+\tfrac{1}{2}v,\, z+\tfrac{1}{2}w)\\ -P(x-\tfrac{1}{2}u,\, y-\tfrac{1}{2}v,\, z-\tfrac{1}{2}w)]; \quad (3)$$

then

$$M_{uvw}(x,y,z)=S_{uvw}(x,y,z)-|D_{uvw}(x,y,z)|. \quad (4)$$

The Fourier series for $S_{uvw}(x,y,z)$ and $D_{uvw}(x,y,z)$ are easily found by substitution in Eqs. (2) and (3) of the Patterson expression,

$$P(x,y,z)=\sum_{hkl}|F_{hkl}|^2\cos 2\pi(hx+ky+lz). \quad (5)$$

Thus,

$$S_{uvw}(x,y,z)=\sum_{hkl}\{|F_{hkl}|^2\cos\pi(hu+kv+lw)\}\\ \times\cos 2\pi(hx+ky+lz), \quad (6)$$

$$D_{uvw}(x,y,z)=\sum_{hkl}\{|F_{hkl}|^2\sin\pi(hu+kv+lw)\}\\ \times\sin 2\pi(hx+ky+lz). \quad (7)$$

The Fourier coefficients in Eqs. (6) and (7) are the entire expressions in brackets and depend on the components of the chosen vector (u,v,w).

It is also possible to express the Fourier coefficients of $M_{uvw}(x,y,z)$ directly in the form,

$$(1/V)\int\exp[-2\pi i(hx+ky+lz)](S-|D|)dxdydz\\ =|F_{hkl}|^2\cos\pi(hu+kv+lw)\\ -(1/V)\int_{(D\geqslant 0)}D\exp[-2\pi i(hx+ky+lz)]dxdydz\\ +(1/V)\int_{(D\leqslant 0)}D\exp[-2\pi i(hx+ky+lz)dxdydz. \quad (8)$$

It is not yet clear that these expressions provide a practical basis for the calculation of Buerger's minimum function. However, the writer has been informed in a private communication from Professor R. Pepinsky that there is a possibility that they will prove adaptable to computation on the Pennsylvania State College high speed computer (X-RAC).[3]

[1] M. J. Buerger: (a) Acta Cryst. 3, 87 (1950); (b) Proc. Natl. Acad. Sci. U. S. 36, 376 (1950); (c) Proc. Natl. Acad. Sci. U. S. 36, 738 (1950); (d) Acta Cryst. 4, 531 (1951).
[2] A. L. Patterson, Phys. Rev. 65, 195 (1944).
[3] R. Pepinsky, J. Appl. Phys. 18, 601 (1947).

R-PAC: Analog Recorder-Playback Computer for Crystal Analysis

RAY PEPINSKY AND PAUL JARMOTZ,
Department of Physics, The Pennsylvania State College
State College, Pennsylvania
(Received January 13, 1953)

R-PAC, the analog recorder-playback computer for crystal analysis, has been discussed briefly previously.[1a] This machine was conceived as an extension of X-RAC and S-FAC and is used in conjunction with these machines.[1b]

Although designed for a variety of purposes, the chief use of R-PAC is for the interpretation of Patterson functions. When applied in the examination of a two-dimensional vector map, an X-RAC signal for two cells of a Patterson function is recorded as an intensity-modulated pattern on photographic film wrapped around a transparent Lucite drum, rotating on its axis and simultaneously displaced in the axial direction. Recording is accomplished by a fast facsimile system. After photographic processing, accomplished by removal of the drum from the rotating mechanism and without removal of the film from the drum, the drum is placed on the playback section, which is separate from the recording section in order to release X-RAC for other computations and in order to simplify various constructional details.

The signal is played back by illuminating the rotating, translating drum from within, and picking up transmitted light with suitably placed optical probes, consisting of small lens systems coupled to photomultipliers. Positioning of the probes, corresponding to a Buerger image polygon,[2] and suitable electronic combination of the resulting photomultiplier signals permit application of any one of a number of image-seeking methods. The final signal is presented as a contour map on a long-persistent-screen oscilloscope, as in the X-RAC presentation.[1c] This contour map can be observed while probes are repositioned in any desired manner and can of course be photographed. The resultant playback signal can also be fed back to the recording machine, in case it is necessary to obtain linearly-undistorted maps or to recombine R-PAC maps subsequently.

When three-dimensional computations are desired, Patterson sections at fixed intervals in a direction perpendicular to the sections are separately computed on X-RAC, recorded on individual drums, and photographically processed. These are then examined by subsequent placement on the R-PAC playback system in suitable combinations, the machine being constructed to accommodate several coaxial drums.

It is a particularly simple matter to compute Buerger's minimum function[2] for image-seeking with this instrument. The individual photomultiplier signals, at a low impedance level and at voltage-division levels, established according to required weighting values, are fed to separate cathodes of a corresponding number of diodes, the plates of which are tied together. A signal corresponding to the minimum of all of the probe photomultiplier voltages is immediately obtained at the common diode plates, since higher signals on the cathodes of the other diodes serve to cut these off. This minimum signal is amplified, contoured, and presented on the persistent-screen CR tube.

T. Noguchi has suggested that when the intensity-modulated maps are scanned by the several optical probes and only the signal which is *the minimum of all those observed at any instant* is passed by the diodes' circuit to the amplifier, nonlinearities between film-density and true Patterson map density can be compensated for in the contour-generating circuit. This is certainly true if the atoms of the density function all have the same or very similar electron-density distribution, and thus the probe outputs are closely similar in weighting. When the electron densities of the atoms differ significantly, the necessity for *different* weightings of the probe outputs (by simple adjustable voltage-dividers at these outputs), in order to apply Buerger's function properly, prevents application of Noguchi's simple nonlinearity compensation scheme. Thus in general a good degree of linearity between the true Patterson map density and the optical probe outputs must be attained. This is the most critical requirement of the R-PAC system, and it must be satisfied both here and in some other applications of this machine. Amplifier circuits must also be accurately normalized, at proper dc and gain levels.

The final model of R-PAC is being constructed with highly linear sweep circuits, similar in accuracy to those utilized in X-RAC and S-FAC. These provide low linear distortion and are particularly important if playback functions are later to be re-

combined via the recorder section. For proper operation of both machines, sweep duty cycles must also be very high, as in present X-RAC and S-FAC circuits; and the quality of the R-PAC circuits must match that of the other machines.

Validity of values of atomic coordinates deduced from the Buerger patterns obtained from the playback machine can be very rapidly examined by calculations on S-FAC, which computes x-ray scattering factors. When a reasonable approximate atomic distribution appears, combined use of X-RAC and S-FAC then permits rapid application of the iterative process for convergence on the complete structure.[3]

R-PAC was originally designed as a device for recording nondistorted X-RAC maps,[1d] for instrumentation of extension of range of X-RAC[1e] and rapid computation of three-dimensional electron density and Patterson maps,[1f] and for application of vector convergence density methods of Patterson map interpretation.[1a,1g] The recording of contour maps alone does not require the *amplitude* linearity essential in general applications of image seeking methods, extension of range, and three-dimensional function computations. R-PAC recording of contour maps will be carried out in all cases where high fidelity of mapping is essential.

The applicability of R-PAC in computations of "difference" functions, and particularly $F_o - F_c$ density maps, in conjunction with both X-RAC and S-FAC is obvious. The machine should be useful as well in examination of Patterson maps from isomorphous crystals containing some atoms of known positions and different densities.

The use of magnetic drum recording and playback, in place of the optical method here presented, has long been considered. The optical procedure directly provides permanent records. A magnetic system, utilizing frequency modulation for signal amplitude fidelity, has many advantages in other respects, however, and *viewing* could be achieved by a simultaneous oscilloscope presentation. It is likely that a final model of the machine will involve both optical *and* magnetic recording-playback schemes, the choice between these, for any computation, depending on the nature of the computation.

The playback machine, with its associated probes, minimizing and contouring circuits, and presentation oscilloscope, is the simplest component of R-PAC, and it can be duplicated in other laboratories, in cases where investigators wish to apply image-seeking methods without journeying to the Pennsylvania State College. The use of X-RAC in preparation of the Patterson functions is of course not essential; but their preparation on X-RAC will certainly be the most rapid technique. The power of the three machines, X-RAC, S-FAC, and R-PAC, in combination does render their integrated use most advisable.

The playback section of R-PAC could also be utilized in the construction of somewhat simplified X-RAC and S-FAC machines, in that the H_1 and 1000 H_2 X-RAC sin and cos waves[1h] could be drum-recorded and utilized to eliminate some sections of the X-RAC frequency generators. This possibility should be considered by anyone desiring to duplicate our machines.

Recognition by Pepinsky of the ease with which Buerger's minimum image-seeking function could be introduced into R-PAC was catalyzed by a private communication from W. J. Taylor of Ohio State University, who had carried out an ingenious mathematical analysis of Buerger's functions, and suggested another procedure for evaluating the minimum function on X-RAC.[4] The entire development has profited from several discussions with A. L. Patterson over the several years during which it has been under consideration. The advice of M. J. Buerger on the computation of his image-seeking functions has also been very useful to us.

The assistance of A. F. Moodie, T. Noguchi, and C. W. Brouse of the School of Chemistry and Physics Shop, The Pennsylvania State College in the mechanical design of a very preliminary model of the playback system is appreciated. Moodie and Noguchi also assisted in some aspects of electronic design of the preliminary model of the playback section. Support of this entire development by the Physics Branch of the U. S. Office of Naval Research is very gratefully acknowledged.

[1] R. Pepinsky: (a) *Computing Methods and the Phase Problem in X-Ray Crystal Analysis* (X-Ray Crystal Analysis Laboratory, The Pennsylvania State College, 1952), pp. 273–274; (b) *ibid.*, pp. 167–318; (c) *ibid.*, pp. 177 and 225–237; (d) *ibid.*, pp. 237, 243; (e) *On the Extension of Range of X-RAC*, Tech. Rept. to Office of Naval Research, Contract N6onr-26916, T. O. 16, November 1, 1949; see also *Computing Methods*, pp. 294–299; (f) *Computing Methods*, pp. 251–253; (g) Proposal to Office of Naval Research, Contract N6onr-26916, T. O. 16, February 6, 1951; (h) *Computing Methods*, p. 169.

[2] M. J. Buerger, Acta Cryst. **4**, 531 (1951).

[3] J. M. Bijvoet, in *Computing Methods* (see reference 1), pp. 10–17.

[4] W. J. Taylor, J. Appl. Phys., **24**, 662 (1953) (see preceding paper).